



Quaestio facti

Revista Internacional sobre Razonamiento Probatorio
International Journal on Evidential Legal Reasoning

2025 | N. 8

ISSN: 2604-6202

ENSAYOS

Fernando Mariath Rechia, *Raciocínio Indutivo no Direito Tributário: inferência e suficiência probatória*

Raúl Díaz Manosalva, *Deliberación, función epistémica y establecimiento de hechos. Bases de una estructura metodológica para el tribunal penal*

Miguel Ángel Zumba Barbosa, *Estándar de prueba en controversias por omisión tributaria: una mirada a la legislación y jurisprudencia ecuatoriana*

María Hernández Moreno, *La prueba pericial científica en manchas de sangre: una reflexión sobre su aplicación en el proceso español. Sesgos y recomendaciones*

Serrana Delgado Manteiga, *La causalidad y la normatividad en la prueba tributaria. Reflexiones a partir de la regla de exclusión probatoria*

CONJETURAS Y REFUTACIONES

Federico Picinali, *Epistemic Injustice in the Criminal Trial: Engaging with Gonzales Rose, Herdy, Jalloh and Owusu-Bempah*

Jacopo Della Torre, *Taking the Evolution of the Standards of Proof for a Criminal Conviction Seriously*

CIENCIA PARA EL PROCESO

Antonio L. Manzanero, Álvaro Alanzavez, Jordi Ferrer Beltrán, Carmen Vázquez, *Falsas creencias sobre la memoria de los testigos en profesionales y estudiantes de ciencias sociales y jurídicas*

Qf

Quaestio facti

Revista Internacional sobre Razonamiento Probatorio
International Journal on Evidential Legal Reasoning

2025 | N. 8

www.quaestiofacti.com

Quedan rigurosamente prohibidas, sin la autorización escrita de los titulares del «Copyright», bajo las sanciones establecidas en las leyes, la reproducción total o parcial de esta obra por cualquier medio o procedimiento, comprendidos la reprografía y el tratamiento informático, y la distribución de ejemplares de ella mediante alquiler o préstamo públicos.

© CÁTEDRA DE CULTURA JURÍDICA

© MARCIAL PONS
EDICIONES JURÍDICAS Y SOCIALES, S. A.
C/ Tamayo y Baus, 7, Bajo izq. – 28004 Madrid
☎ (91) 304 33 03
www.marcialpons.es
ISSN: 2604-6202
Depósito legal: M. 3.518–2020
Fotocomposición: Ene Estudio
Impresión: ELECÉ, Industria Gráfica, S. L.

DIRECTORES

Diego Dei Vecchi
Universitat de Girona, España

Jordi Ferrer Beltrán
Universitat de Girona, España

COMITÉ DE REDACCIÓN

Daniela Accatino
Universidad Austral de Chile, Chile

Daniel González Lagier
Universidad de Alicante, España

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

Carmen Vázquez
Universitat de Girona, España

SECRETARÍA DE REDACCIÓN Y EDICIÓN

Anna de Castro Cros
Universitat de Girona, España

COMITÉ EDITORIAL

Christian Dahlman
Lund University, Suecia

Antonio Manzanero
Universidad Complutense de Madrid, España

Mauricio Duce
Universidad Diego Portales, Chile

Paul Roberts
University of Nottingham, Reino Unido

Mercedes Fernández
Universidad de Alicante, España

Sarah Summers
University of Zurich, Suiza

Raymundo Gama
Instituto Tecnológico Autónomo de México

Giovanni Tuzet
Università Bocconi, Italia

Ho Hock Lai
National University of Singapore, Singapur

Jonatan Valenzuela
Universidad de Chile, Chile

CONSEJO ASESOR

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

ÍNDICE

ENSAYOS

Fernando Mariath Rechia, <i>Raciocínio Indutivo no Direito Tributário: inferência e suficiência probatória</i>	11
Raúl Díaz Manosalva, <i>Deliberación, función epistémica y establecimiento de hechos. Bases de una estructura metodológica para el tribunal penal</i>	25
Miguel Ángel Zumba Barbosa, <i>Estándar de prueba en controversias por omisión tributaria: una mirada a la legislación y jurisprudencia ecuatoriana</i> ...	57
María Hernández Moreno, <i>La prueba pericial científica en manchas de sangre: una reflexión sobre su aplicación en el proceso español. Sesgos y recomendaciones</i>	89
Serrana Delgado Manteiga, <i>La causalidad y la normatividad en la prueba tributaria. Reflexiones a partir de la regla de exclusión probatoria</i>	109

CONJETURAS Y REFUTACIONES

Federico Picinali, <i>Epistemic Injustice in the Criminal Trial: Engaging with Gonzales Rose, Herdy, Jalloh and Owusu-Bempah</i>	143
Jacopo Della Torre, <i>Taking the Evolution of the Standards of Proof for a Criminal Conviction Seriously</i>	155

CIENCIA
PARA EL PROCESO

Antonio L. Manzanero, Álvaro Alanzavez, Jordi Ferrer Beltrán, Carmen Vázquez, <i>Falsas creencias sobre la memoria de los testigos en profesionales y estudiantes de ciencias sociales y jurídicas.....</i>	219
---	-----

ENSAYOS

RACIOCÍNIO INDUTIVO NO DIREITO TRIBUTÁRIO: INFERÊNCIA E SUFICIÊNCIA PROBATÓRIA

Fernando Mariath Rechia
Doutorando em Direito (USP)
Juiz Federal em São Paulo
fmrechia@usp.br

RESUMO: Este artigo tem o objetivo de analisar o raciocínio indutivo e a sua aplicação para a resolução de questões fáticas em matéria tributária. Para tanto são apresentadas três espécies de indução (generalizadora, causal e analógica), examinados os seus usos (no Direito Tributário e fora dele) e a sua força (inclusive as suas limitações). O artigo conclui que a qualidade do raciocínio probatório é influenciada por duas variáveis. De um lado, pela observância a critérios relativos à formação e avaliação dos elementos probatórios que servem de base à premissa fática da sentença, tais como a grandeza, variedade e representatividade dos elementos de prova. De outro, pelo *standard* de prova aplicável ao caso: quanto mais rigoroso for o grau de suficiência exigido para que determinada hipótese probatória seja considerada provada, mais forte deve ser o grau de segurança inferencial.

PALAVRAS-CHAVE: prova; lógica; epistemologia jurídica; inferência.

INDUCTIVE REASONING IN TAX LAW: INFERENCE AND EVIDENTIARY SUFFICIENCY

ABSTRACT: This article aims to analyze inductive reasoning and its application to resolve factual issues in tax matters. To this end, three types of induction are presented (generalization, causal and analogical), as well as their uses (in Tax Law and outside of it) and their strength (including their limitations). The article concludes that the quality of the evidentiary judgment is influenced by

two variables. On the one hand, by meeting the criteria regarding the formation and evaluation of evidence that serve as a basis for the sentence, such as the magnitude, variety and representativeness of the evidence. On the other hand, by the standard of proof applicable to the case: the stricter the degree of sufficiency required for certain hypotheses to be considered proven, the stronger must be the degree of inferential security.

KEYWORDS: evidence; logic; legal epistemology; inference.

SUMARIO: 1. INTRODUÇÃO.— 2. ESPÉCIES DE RACIOCÍNIO INDUTIVO: REGRAS DE INFERÊNCIA E OS SEUS USOS NO DIREITO: 2.1. Indução Generalizadora (“Generalização”); 2.2. Indução Causal; 2.3. Indução Analógica (“Analogia”).— 3. SUFICIÊNCIA PROBATÓRIA E STANDARDS DE PROVA.— 4. CONCLUSÃO.— REFERÊNCIAS

1. INTRODUÇÃO

Um determinado conjunto de empresas é autuado sob alegação de fraude. De acordo com o Fisco, embora essas sociedades se apresentem como autônomas, na verdade elas constituem uma única empresa. Isto é, embora formalmente separadas, individualmente não possuem autonomia, de modo que onde se dizia haver oito empresas, há apenas uma. Em função disso, o regime fiscal não poderia ser definido pela receita bruta individual de cada sociedade, mas do somatório de todas, circunstância que conduz à exclusão das autuadas do regime simplificado de tributação (“Simples Nacional”).

Ao contrário de parte relevante das questões tributárias, o acerto dessa hipotética atuação não pode ser avaliado apenas pelo estudo dos conceitos de “fraude”, “autonomia empresarial” ou “receita bruta” (*quaestio juris*). É preciso examinar os fatos (ou os enunciados fáticos) a fim de responder perguntas como: as empresas têm sedes diferentes ou estão no mesmo estabelecimento? Cada empresa possui colaboradores e meios suficientes à produção/prestação do serviço (ex. maquinário) ou esses elementos são compartilhados? O quadro de sócios é idêntico? As empresas prestam serviços a terceiros ou apenas entre si? Mais do que isso, é preciso avaliar quais elementos de prova (e em que intensidade) devem estar presentes para autorizar a conclusão de que a fraude está provada no caso concreto (*quaestio facti*). O percurso para a resolução dessas questões —de enorme complexidade, mas indispensáveis à solução do caso— somente pode ser percorrido por meio de um raciocínio probatório.

Um raciocínio é uma operação intelectual por meio da qual se chega a uma conclusão a partir de determinadas premissas. Trata-se de uma atividade que é realizada a todo o tempo, seja para a tomada de decisões banais, como a necessidade de levar um guarda-chuva diante de uma previsão de chuva, até decisões complexas, como a avaliação sobre a suficiência de determinados meios de prova para confirmar um auto de infração fiscal ou a culpa de um réu em uma ação penal.

Há basicamente dois tipos: o raciocínio dedutivo e o indutivo.

O raciocínio dedutivo é aquele por meio do qual a passagem das premissas à conclusão se dá por implicação lógica: a conclusão decorre necessariamente das premissas. Não há como as premissas serem verdadeiras e a conclusão obtida por dedução ser falsa. O raciocínio dedutivo é capaz de fornecer segurança inferencial absoluta: as suas premissas fornecem suporte definitivo à conclusão. O critério de verificação aplicável ao raciocínio dedutivo é o da validade (Kelley, 2014, p. 83-84).

Em um exemplo singelo: imagine-se que há uma caixa repleta de laranjas. A primeira premissa (geral) do raciocínio é a de que todas as laranjas dessa caixa estão doces. A segunda é a de que uma laranja foi retirada dessa caixa. A conclusão, obtida por dedução, é de que essa laranja é *necessariamente* doce.

O raciocínio indutivo, por outro lado, é aquele por meio do qual a passagem das premissas à conclusão se dá por probabilidade. Se as premissas são verdadeiras, a conclusão é provavelmente verdadeira. O raciocínio indutivo é capaz de fornecer segurança gradual: as suas premissas fornecem suporte *parcial* à conclusão. O critério de verificação aplicável ao raciocínio dedutivo é o da força.

Utilizando-se a mesma caixa de laranjas, no caso do raciocínio indutivo a premissa é a de que dez laranjas retiradas da caixa estavam doces. A segunda é a de que uma nova laranja foi retirada dessa caixa. A conclusão, obtida por indução, é de que essa laranja é *provavelmente* doce.

Na primeira hipótese, como já adiantado, é possível afirmar sem risco de erro a qualidade do exemplar de laranja retirada da caixa. Na segunda há sempre a chance de que a laranja esteja azeda. A incerteza demonstrada no segundo exemplo conduz à seguinte indagação: por que utilizar o argumento indutivo, e não o dedutivo, se aquele não é capaz de proporcionar certeza? São duas as razões.

De um lado, pois a dedução não é capaz de incorporar novos conhecimentos àqueles já concentrados na premissa (característica da “reflexividade” dos raciocínios dedutivos), apenas de organizá-los de forma mais clara. Isto é, a partir do momento em que se sabe que toda a classe alcançada pela conclusão possui uma determinada característica (todas as laranjas da caixa são doces), a conclusão de que uma de suas espécies possui essa característica (a laranja retirada da caixa é doce) é absolutamente trivial.

De outro lado, uma conclusão obtida a partir do raciocínio dedutivo depende da existência de conhecimentos (premissas) muito abrangentes, que nem sempre —na verdade, quase nunca— estão disponíveis. Nem sempre alguém terá condições de tomar como dado a informação de que “todas as laranjas da caixa são doces”. Há situações em que a caixa é muito grande, contém muitas laranjas ou em que o acesso ao seu conteúdo é apenas parcial.

Ao contrário do que se poderia pensar, a relação entre os raciocínios dedutivo e indutivo não é de oposição, mas de complementaridade (Kelley, 2014, p. 85). Um mesmo enunciado pode exercer a função de premissa em um determinado raciocínio e de conclusão em outro. No exemplo das laranjas, o enunciado “todas as laranjas da

caixa estão doces” pode ser a premissa de um raciocínio dedutivo (cuja conclusão é a de que uma laranja retirada da caixa será *necessariamente* doce) ou a conclusão de um raciocínio indutivo (obtida por meio de uma generalização da avaliação de um determinado número de laranjas contidas na caixa). Os enunciados não vêm com rótulos fixos de “premissa” ou “conclusão”.

Os raciocínios dedutivo e indutivo, portanto, possuem funções diferentes e estão disponíveis em situações diversas. No âmbito do Direito, cada um deles desempenha um papel na justificação — “interna” e “externa” — das decisões. A dedução é utilizada para testar a racionalidade formal da decisão, a fim de verificar se ela é *internamente* justificada (ou seja, se a solução decorre logicamente das premissas apresentadas pelo julgador) (Wróblewski, 1979, p. 63; Guastini, 2011, p. 257). Já a indução consiste em um teste de racionalidade *substancial* da decisão: analisa a correção das premissas, em especial da premissa probatória (Chiassoni, 2007, p. 14-15).

Logo, em todas aquelas situações nas quais é preciso obter informações até então desconhecidas, a utilização do raciocínio indutivo em detrimento do dedutivo não é uma opção, mas uma necessidade.

Como visto, contudo, a indução jamais será capaz de fornecer segurança inferencial absoluta, somente relativa: em maior ou menor grau. Como afirmado por Gascón Abellán (2010, p. 61), “*um conhecimento provável segue sendo um conhecimento objetivo, pois a falta de certeza absoluta não pode conduzir a ignorar que existem diferentes graus de certeza*”. Daí a sua relação com o tema dos *standards* de prova, instrumentos responsáveis por fixar o grau de suficiência exigido para que determinada hipótese seja considerada provada. Quanto mais exigente o parâmetro, maior a segurança exigida em relação ao raciocínio inferencial.

Daí que se faz necessário analisar (i) as regras aplicáveis ao raciocínio indutivo, a fim de viabilizar a construção de inferências dotadas de maior solidez/segurança, e (ii) os *standards* de prova, com a finalidade de avaliar se a inferência é suficiente para comprovar a hipótese probatória.

2. ESPÉCIES DE RACIOCÍNIO INDUTIVO: REGRAS DE INFERÊNCIA E OS SEUS USOS NO DIREITO

A estrutura do raciocínio indutivo admite diferentes usos que permitem atingir variadas finalidades. Daí se falar em diversas espécies de raciocínios ou técnicas que se valem da lógica indutiva. São três as espécies com maior utilidade para o Direito: indução generalizadora, indução causal e indução analógica.

2.1. Indução Generalizadora (“Generalização”)

A indução generalizadora é a técnica por meio da qual é possível obter uma conclusão a respeito de uma classe (“população”) a partir da avaliação de um número limitado de suas espécies (“amostra”) (Rudinow e Barry, 2007, p. 231). Em outras palavras, em razão de os exemplares que integram a amostra possuírem certas características comuns, conclui-se que todos os membros daquele grupo possuem essa característica.

É o que se verifica no exemplo apresentado acima a respeito das laranjas retiradas de uma caixa. A avaliação em relação a um conjunto limitado de laranjas (“dez laranjas que eu provei dessa caixa estavam doces”) é expandida para englobar outros exemplares da fruta contidos na caixa, ainda que estes não tenham sido avaliados diretamente. O mesmo ocorre quando se conclui que todos os corvos são pretos ou que todos os metais dilatam com o calor, ainda que nem todos os exemplares que integram as classes “corvos” e “metais” tenham sido individualmente examinados/testados.

Esse tipo de raciocínio é pervasivo na experiência humana. Trata-se de um instrumento sem o qual não seria possível organizar o conhecimento – em rigor, a sua própria obtenção ficaria seriamente prejudicada. A noção de “experiência” teria de ser revista, na medida em que cada ocorrência seria encarada como única. As assim chamadas “máximas de experiência” com as quais evitamos colocar o dedo em tomadas ou em painéis ferventes (apenas para ficar em duas de suas utilizações mais singelas) seriam impossibilitadas, na medida em que elas são formadas a partir de padrões e regularidades.

O uso da generalização é particularmente relevante no âmbito do Direito, em que o legislador se vale da utilização de termos gerais para atingir um objetivo bastante específico: regular conjuntos de condutas atuais e futuras. Por conseguinte, o material bruto a partir do qual as normas são reconstruídas é necessariamente fruto de um processo de generalização (Schauer, 1991, p. 17-111; Larenz, 1997, p. 294-295). Mais do que uma contingência, esse processo representa a concretização dos princípios do Estado de Direito, Legalidade e Igualdade, que exigem, entre outras coisas, que os enunciados normativos não se refiram a indivíduos ou situações particulares, mas a conjuntos e situações abstratamente consideradas. Assim, para que as leis sejam gerais e abstratas, elas devem se destinar a classes, e não a particulares (Schauer, 2003, p. 4). A título de exemplo, a exigência de que somente pessoas com mais de 35 anos possam ocupar os cargos de Presidente da República, Senador e Ministro do Supremo Tribunal Federal é fruto de um raciocínio indutivo, precisamente o de que pessoas com menos idade não têm —como regra *geral*— a experiência de vida e a maturidade necessárias ao exercício dessas funções. O mesmo ocorre quando o legislador infere que proprietários de imóveis localizados em áreas nobres e com maior metragem possuem maior capacidade contributiva para efeito de recolhimento de IPTU. Em ambos os casos é possível imaginar um sem-número de exceções à generalização, circunstância que apenas demonstra a natureza probabilística —e não absoluta— do raciocínio indutivo.

A qualidade da generalização é dependente da qualidade dos critérios utilizados para formar os elementos que servem de base à conclusão/generalização. Em outras palavras, a segurança da conclusão obtida por meio desta primeira espécie de raciocínio indutivo varia conforme a observância das regras aplicáveis à formação da “amostra” (Soccio e Barry, 1998, p. 265-266; Rudinow e Barry, 2007, p. 232-233).

Em primeiro lugar, o critério do *tamanho* da amostra exige que a quantidade de espécies examinadas seja numericamente proporcional à classe a respeito da qual a generalização se refere. Retomando o exemplo já apresentado, seria insuficiente a amostra que, para concluir a respeito da qualidade de duzentas laranjas armazenadas em uma caixa, avaliasse apenas dois dos seus exemplares. Da mesma forma, atirar uma moeda para cima por cinco vezes —e constatar que em quatro delas o resultado foi “coroa”— configura amostra insuficiente para concluir que toda vez que alguém arremessar uma moeda para cima há apenas 20% de chance de dar “cara”.

Em segundo lugar, o critério da *representatividade* exige que a amostra reflita adequadamente características comuns à classe (e não a uma porção dela). Daí que mesmo a avaliação de um número considerável de laranjas (tamanho da amostra) não seria suficiente caso todos esses exemplares tivessem sido retirados da parte de cima da caixa. Isso porque, é possível que as laranjas de pior qualidade —propositalmente ou não— tenham sido inseridas na parte de baixo da caixa, de tal forma que a avaliação que leva em conta apenas as laranjas posicionadas na parte de cima não seria representativa do total.

Em terceiro lugar, o critério da *aleatoriedade*, no sentido de que cada espécie da classe (grupo objeto da generalização) deve ter a mesma chance de integrar a amostra. No exemplo acima é fácil perceber que ao limitar a seleção de exemplares às laranjas situadas na parte de cima da caixa, a seleção da amostra não atendeu à exigência de aleatoriedade.

Em se tratando da indução generalizadora em raciocínios probatórios, a “amostra” é composta pelo conjunto de elementos considerados pelo juízo para suportar a hipótese que ocupará a posição de premissa menor/fática da decisão. No exemplo que deu início a este trabalho, a amostra seria integrada pelas respostas às indagações formuladas a respeito das oito empresas autuadas pela fiscalização tributária. Isto é, os elementos de prova utilizados para decidir se elas possuíam autonomia ou não passavam de um artifício fraudulento para permitir o enquadramento em regime simplificado de tributação. A formação desse conjunto que será a base da tomada de decisão probatória deverá respeitar os critérios de tamanho/grandeza, representatividade e aleatoriedade, a fim de assegurar a força da inferência.

O caráter quase onipresente da indução generalizadora em raciocínios probatórios acaba por tornar frequentes os casos em que o seu uso não atende aos critérios apresentados acima. Em um caso julgado pelo Conselho Administrativo de Recursos Fiscais (CARF), uma empresa distribuidora havia sido autuada com aplicação de multa qualificada pela Receita Federal do Brasil sob a alegação de simulação na forma de contratação dos representantes comerciais que lhe prestavam serviço. De

acordo com a fiscalização, a autuada criou artificialmente pessoas jurídicas para ocultar a relação de emprego mantida com cada representante. A imputação da fraude foi derrubada pelo CARF em virtude de problemas no raciocínio probatório empregado, notadamente no que se refere à generalização apressada que acabou por desprezar os parâmetros de tamanho, representatividade e aleatoriedade na formação e avaliação dos elementos de prova. Os principais problemas identificados foram dois. De um lado, pois a conclusão sobre a ilicitude na contratação de 441 representantes comerciais foi amparada em um conjunto reduzido de elementos: provas orais produzidas pelas partes em poucas reclamações trabalhistas. De outro, pois a fiscalização deixou de incluir nesse conjunto de análise os elementos de prova oriundos de reclamações trabalhistas transitadas em julgado declarando a inexistência de vínculo empregatício entre a autuada e ao menos 25 de seus representantes comerciais. O seguinte trecho do voto vencedor é elucidativo: “[a]s ilações da fiscalização sobre a existência de fraude estão fortemente apoiadas em extração de dados a partir de prova oral produzida pelas partes em reclamações trabalhistas, num exercício de generalização de depoimentos e testemunhos, já que baseado em casos específicos do grupo de prestadores de serviços, procedimento que pede cautela e equilíbrio” (Conselho Administrativo de Recursos Fiscais, Acórdão n. 2401005.952, de 27 de janeiro de 2019).

2.2. Indução Causal

A indução causal é a técnica utilizada para inferir relações de causa e consequência entre dois eventos (Soccio e Barry, 1998, p. 331).

Abstraído-se a complexa discussão a respeito do que configura uma “causa” (se necessária, suficiente ou meramente contributiva) e dos diversos critérios criados para aferir a relação de causalidade entre dois eventos (os quais remontam ao século XIX, com John Stuart Mill (2011, p. 450-479), essa espécie de raciocínio indutivo pode ser utilizada para cumprir diversas finalidades.

De um lado, é possível empregar a indução causal para influenciar comportamentos, como nos seguintes exemplos: “pare de fumar, você vai ter câncer de pulmão”; “se você não estudar, não será bem-sucedido”; “não roube, você vai ser condenado a uma pena de quatro a dez anos de reclusão”.

De outro, é possível utilizá-la como instrumento para elucidação de fatos. Trata-se da lógica por trás dos indícios e das presunções (Godden e Walton, 2007, p. 313; Machado Segundo e Machado, 2014, p. 159). Isto é, a relação causal entre dois eventos é utilizada para concluir a respeito da ocorrência de um evento a partir da prova de um evento diverso que, com ele, guarda uma relação de causa e efeito (Hurley e Watson, 2018, p. 37; Rudinow e Barry, 2007, p. 230; Kelley, 2014, p. 415). É o que se verifica na costumeira relação que se estabelece entre a fumaça e o fogo, no sentido de que, onde há o primeiro, há o segundo. Nesse caso, a existência do “fogo” é inferida a partir da prova da fumaça e da justificação da relação de probabilidade

causal entre as duas ocorrências. Dessa forma, ainda que não haja nenhuma filmagem ou testemunha da existência do fogo (prova direta), a sua ocorrência é provada de forma indireta.

Trata-se de raciocínio frequentemente utilizado no Direito. A própria legislação possui exemplos da sua aplicação, como é o caso do artigo 42 da Lei n. 9.430, segundo o qual a omissão de rendimentos é presumida diante da existência de depósitos bancários não declarados. Da mesma forma, a indução causal é empregada em raciocínios judiciais, em especial quando há necessidade de analisar comportamentos e finalidades. Dada a impossibilidade de ter acesso imediato/direto à mente humana, a consciência e vontade do agente (dolo), por exemplo, será sempre objeto de prova indireta.

A grande dificuldade da indução causal reside na confusão comumente estabelecida entre “correlação” e “causalidade”. Isso porque, o simples fato de duas coisas andarem juntas não significa que uma cause a outra. Algumas ocorrências curiosas são suficientes para provar o argumento: apurou-se que entre os anos de 2012 a 2021 as buscas pelo termo “zumbis” no *google* foram proporcionais ao número de corretores de imóveis no Estado de North Dakota; entre os anos de 2000 e 2009, o consumo per capita de margarina acompanhou o percentual de divórcios no estado do Maine; por fim, entre os anos de 1999 e 2009, o número de filmes que contaram com a participação do ator Nicolas Cage foi proporcional ao número de pessoas mortas por afogamentos em piscinas nos Estados Unidos¹. Contudo, independentemente da avaliação que se faça em relação à qualidade dos filmes protagonizados pelo mencionado ator (ou das pesquisas envolvendo zumbis ou margarinas), não há nenhuma relação causal entre as duas ocorrências —muito embora a evolução dos números demonstre uma correlação, no sentido de que o aumento de uma é seguido do aumento proporcional da outra, e vice-versa.

2.3. Indução Analógica (“Analogia”)

A indução analógica é a técnica por meio da qual determinadas semelhanças entre dois objetos ou eventos são utilizadas para inferir semelhanças adicionais em relação a ambos, circunstância que justificaria a identidade —ou diferença— de tratamento. Como é possível perceber, o argumento analógico guarda inúmeras semelhanças com a indução generalizadora, razão pela qual alguns autores caracterizam a analogia como uma forma de generalização (Soccio e Barry, 1998. p. 314-315).

No Direito, a analogia costuma ser lembrada quase que exclusivamente como método para colmatar lacunas, situação caracterizada quando uma determinada classe de fatos juridicamente relevantes não é regulada por nenhuma norma (explícita) (Hage, 1997, p. 118 e 120; Chiassoni, 2007, p. 196 e ss.). Essa aplicação do raciocí-

¹ Disponível em: <https://www.tylervigen.com/spurious-correlations>. Acesso em: 11.03.2024.

nio analógico possui função normativa, e não probatória: ela interfere na construção da premissa maior do raciocínio decisório. Daí se dizer que a analogia é uma técnica de construção jurídica (Larenz, 1997, p. 520; Guastini, 2015, p. 20; Villa, 2012, p. 35; Chiassoni, 2007, p. 49-50; Falcão, 1959, p. 84-85; Torres, 2006, p. 33). O seu emprego permite atribuir a um caso não regulado a mesma consequência jurídica aplicável a um (segundo) caso regulado, sob o argumento de que a semelhança entre ambos exigiria a mesma disciplina jurídica (Bobbio, 1993, p. 266; Torres, 1988, p. 262). Um exemplo da sua utilização se deu no julgamento da ADI 4277 (STF, n. 4.277, de 05 de maio de 2011), quando o Supremo Tribunal Federal, de forma unânime, estendeu o conceito de união estável para englobar relações entre pessoas do mesmo sexo (artigo 226, § 3º, da Constituição). A despeito de o texto constitucional limitar a chancela estatal às uniões “entre o homem e a mulher como entidade familiar”, a decisão reconheceu que essa diferença é irrelevante sob a perspectiva dos requisitos necessários à configuração do instituto da união estável: “convivência contínua e pública direcionada para a formação de uma autônoma unidade doméstica/familiar”². Sob esses aspectos, concluiu ser injustificado o tratamento distinto às uniões homoafetivas. O julgamento, contudo, teria resultado diverso se, no lugar de regra que disciplina a união estável, estivesse em análise uma norma que cria ou aumenta um tributo (Schoueri, 2015, p. 759). Isso porque, o Direito brasileiro possui regras constitucionais (artigo 150, I, da Constituição Federal de 1988) e legais (artigo 108, §1º, do Código Tributário Nacional) que proíbem a utilização da técnica da analogia para criar um aumentar tributo (vedação à utilização da analogia gravosa com função normativa).

Contudo, a função normativa não esgota a aplicação do raciocínio por indução analógica no Direito. Assim como verificado em relação à generalização, a analogia possui grande utilidade ao raciocínio probatório. A partir do conhecimento sobre as características do objeto “A” e da constatação da sua semelhança em relação a alguma propriedade tida como essencial ou relevante do objeto “B”, é possível inferir que essa semelhança alcança outras características que o sujeito desconhece em relação a “B”. Nessa hipótese, a analogia não atua para estender o campo de aplicação da norma (construção jurídica), mas para qualificar fatos de acordo com o seu campo de referência (interpretação “em concreto”). Daí que a sua utilização é autorizada mesmo em matéria tributária. Um exemplo basta para demonstrá-lo. Imagine-se um caso envolvendo o Imposto sobre Transmissão de Bens Imóveis (ITBI), cuja base de cálculo corresponde ao valor de mercado do imóvel. João, adquirente de apartamento localizado no Prédio “Carlos Drummond”, apresenta declaração tributária informando que o valor venal do imóvel é de R\$ 100.000,00. O Fisco municipal, suspeitando que o valor apresentado tenha sido objeto de simulação, procede à investigação em relação aos valores de outros imóveis localizados no mesmo edifício, identificando dois apartamentos no mesmo andar do prédio adquirido por João, anunciados à

² Trecho extraído do voto condutor, proferido pelo Min. Ayres Britto (Relator).

venda no mesmo período, pelo montante de R\$ 400.000,00. Nesse caso, embora o Fisco não tenha acesso aos valores do negócio jurídico que resultou na transmissão do apartamento a João (que pode inclusive ter pagado uma parcela do montante em dinheiro para evitar a tributação), a semelhança entre as características deste imóvel com os outros dois (mesmo prédio, andar e período de anúncio/venda) permite inferir que os valores declarados são inferiores àqueles efetivamente praticados.

Como é possível perceber, a força do raciocínio de indução analógica está atrelada à semelhança entre os objetos ou eventos comparados. Os principais critérios para avaliá-la são os seguintes: relevância da similaridade (deve ter por objeto características essenciais, e não secundárias); número de semelhanças (quanto maior, mais força terá o raciocínio analógico); natureza e grau das diferenças (as diferenças entre os dois objetos ou eventos comparados devem ser mínimas ou dizer respeito a características não essenciais à finalidade da comparação) (Hurley e Watson, 2018, p. 541-543).

O principal defeito no raciocínio analógico é caracterizado como um problema de relevância, dando lugar à chamada “falácia da analogia”. Esse erro na utilização da técnica ocorre quando são ignoradas distinções relevantes entre os objetos e eventos comparados, ao mesmo tempo em que a comparação é levada adiante a partir de semelhanças em relação a características irrelevantes ou não essenciais (Damer, 2013, p. 162). Imagine-se o exemplo de um indivíduo que busca substituir o seu carro *Mercedes*, de cor prata, sob a seguinte exigência: a de que o novo veículo, assim como o seu atual, lhe cause pouca “dor de cabeça”. Num primeiro cenário hipotético, esse indivíduo pensa em comprar um modelo diferente do seu, mas da mesma marca alemã, sob o argumento de que uma série de procedimentos e peças da montadora são compartilhados entre os dois veículos, o que permite concluir que há uma maior probabilidade de que o novo veículo, ainda que de um modelo diferente, não lhe trará dores de cabeça. Num segundo cenário, esse indivíduo formula o seguinte raciocínio: “o meu veículo atual tem a cor prata e não me causa problemas, logo, se eu adquirir um novo veículo, seja de qual marca for, mas de cor prata, ele também não me causará dores de cabeça”. Examinadas as duas variações é fácil concluir que o segundo raciocínio incorre na mencionada falácia da analogia, haja vista que a identidade entre as cores dos carros, diferentemente da identidade de fabricantes, é uma característica meramente acidental que nada diz a respeito da qualidade das peças, da produção e da durabilidade dos veículos.

No exemplo sobre a apuração do ITBI na venda de imóvel, a força do raciocínio analógico seria drasticamente reduzida caso ao invés de dois apartamentos, a comparação tivesse por objeto apenas um; se ao invés de localizado no mesmo andar, o apartamento anunciado estivesse localizado em andar bem mais alto; e, por fim, se ao contrário de o anúncio ter sido realizado no mesmo mês da aquisição do imóvel pelo contribuinte, tivesse sido feito mais de um ano depois. Tal como verificado acima em relação ao exemplo do automóvel, cada uma das diferenças em relação ao imóvel que foi objeto da declaração tributária fragiliza a inferência promovida pela fiscalização municipal.

3. SUFICIÊNCIA PROBATÓRIA E STANDARDS DE PROVA

A análise dos critérios a serem observados para maximizar a força da inferência probatória representa apenas parte do estudo sobre o raciocínio indutivo no Direito. Isso porque, a força que a inferência possui nada diz a respeito da força que ela *deveria ter* para atingir determinada finalidade (provar que houve simulação na declaração do valor de venda do imóvel para efeito de recolhimento do imposto sobre transmissão de propriedade, por exemplo).

A questão elementar a ser investigada, portanto, pode ser resumida em uma palavra: *suficiência*. Isto é, as premissas fornecem suporte *suficiente* à conclusão? Transpondo a indagação ao direito probatório: os meios de prova (aquilo que é utilizado para provar) são *suficientes* para comprovar a hipótese probatória (aquilo que deve ser provado)?

A resposta a essa pergunta é influenciada por duas variáveis.

De um lado, mais óbvio, pela força das premissas. Premissas mais fortes possuem maior aptidão para suportar conclusões. Ou seja, para concluir a respeito da qualidade das laranjas inseridas numa caixa, a premissa que avaliou 80% dos seus exemplares é mais forte do que aquela que avaliou apenas 20%.

De outro lado, contudo, a suficiência também pode ser aferida pela margem de erro admitida pela hipótese/conclusão. Hipóteses que admitem maior margem de erro são mais facilmente suportadas por premissas —e vice-versa. Imagine-se o seguinte exemplo: um indivíduo vai a um restaurante, come peixe e horas após tem uma reação alérgica; uma semana depois, o mesmo indivíduo vai a um outro restaurante, come peixe e em seguida é acometido por uma nova reação alérgica. A partir dessas ocorrências (premissas) é possível formular duas hipóteses: a primeira, de que esse indivíduo é alérgico a peixe; a segunda, de que há *uma chance* de que esse indivíduo seja alérgico a peixe. Como é possível perceber, a primeira hipótese possui uma margem de erro bem mais restrita quando comparada à segunda. Dadas as duas ocorrências examinadas, é possível afirmar que há provas suficientes de que esse indivíduo “tem uma chance” de ser alérgico a peixe. Por outro lado, essas mesmas ocorrências são insuficientes para concluir que ele “é” (ou que ele muito provavelmente é) alérgico a peixe, hipótese cuja comprovação exigiria a busca por novos e melhores meios de prova (possivelmente a realização de um teste alergênico para a proteína do peixe). Nesse momento surge a seguinte indagação: deve esse indivíduo orientar a sua conduta a partir de uma hipótese pelo simples fato de que ela está “suficientemente suportada” pelas premissas de que ele dispõe?

A pergunta conduz a uma conclusão importante: a margem de erro admitida à hipótese/conclusão não é dada pela força das premissas, tampouco pela avaliação da suficiência atual das premissas para suportar a hipótese. Na verdade, o que a define são as consequências a serem suportadas em caso de erro. Retomando o exemplo:

a consequência verificada ao se admitir a hipótese mais rigorosa —de que aquele indivíduo é alérgico a peixe— é não comer peixe até que sejam obtidas melhores evidências que a confirmem ou infirmem. Caso a hipótese seja refutada —ou seja, caso o teste resulte negativo para alergia a peixe—, a consequência do erro terá sido a de que esse indivíduo foi privado à toa do consumo dessa proteína. De outro lado, contudo, caso esse mesmo indivíduo seguisse comendo peixe, a consequência —caso ele fosse, de fato, alérgico— poderia ser um novo episódio alérgico, de maior gravidade, que poderia resultar na sua hospitalização. Logo, ao confrontar os dois cenários, a gravidade da consequência impõe a adoção da hipótese com a margem de erro mais estreita para orientar a conduta, ainda que isso signifique a necessidade de buscar novos e melhores meios de prova para que ela seja suficientemente comprovada (ou refutada).

No Direito, por sua vez, é o próprio ordenamento jurídico que estabelece, expressa ou implicitamente (geralmente nesta segunda forma), a margem de erro admitida às hipóteses probatórias/conclusões. Essa definição é realizada por meio de um instituto cujo nome —assim como os principais desenvolvimentos doutrinários— é herado do direito anglo-saxão: os *standards* de prova. Esses *standards* nada mais são do que índices de *suficiência* probatória. Em outras palavras, graus de corroboração que o Direito exige para que uma determinada hipótese seja considerada como provada (Ferrer Beltrán, 2021, p. 29; Hellström, 2009, p. 1; Ávila, 2018, p. 117). Em uma linguagem figurativa, eles são como as barras horizontais (“sarrafo”) que um atleta de salto em altura deve transpor para vencer a prova.

De forma semelhante ao observado no exemplo da “alergia a peixe”, a calibragem dos *standards* de prova é baseada em ampla medida nas consequências enfrentadas em caso de erro na avaliação da hipótese. Daí que, a depender da natureza da infração e da gravidade das sanções, entre outros fatores, mais exigente será o *standard* aplicável ao caso.

No caso de um processo criminal, por exemplo, o desfecho —ao menos de forma ideal— será um dos seguintes: condenação ou absolvição. A consequência em caso de erro na avaliação da hipótese será a condenação de um inocente ou a absolvição de um culpado. O Direito brasileiro —nesse ponto alinhado aos principais sistemas jurídicos globais— reputa como mais grave a primeira consequência quando comparada à segunda. Não por acaso há uma presunção constitucional de inocência em favor do réu. Esses pressupostos conduzem à fixação de um *standard* de prova mais rigoroso à hipótese acusatória, de tal forma que a sua preponderância pressupõe o atingimento do elevado parâmetro da prova acima de qualquer dúvida razoável. Essa exigência produz a consequência inevitável de aumentar o número de absolvições de indivíduos que deveriam ser condenados. Por outro lado, ela produz o efeito almejado de reduzir o cômputo de condenações de réus inocentes.

Em sentido inverso, no caso de uma demanda cível entre dois particulares envolvendo a cobrança de determinados valores, a consequência será patrimonial. Em caso de erro, haverá condenação para exigir o pagamento de montante que não era devido

ou, alternativamente, deixará de haver o pagamento de quantia que era efetivamente devida. Em ambas as situações, a própria lógica da disputa não permite —abstraídas discussões envolvendo relações de consumo e demais disputas compostas por parte hipossuficiente— a fixação de um *standard* de prova mais elevado para alguma das partes. Nesse caso, ao contrário do exemplo anterior, o sucesso da demanda —ao menos sob a perspectiva da premissa probatória da decisão— será daquele que atingir o parâmetro da preponderância da prova, ou seja, de que há mais probabilidade de que a sua hipótese seja verdadeira do que o contrário.

4. CONCLUSÃO

Conforme examinado ao longo deste breve ensaio, o raciocínio indutivo é o meio pelo qual o raciocínio probatório é levado adiante. Ele mantém relação com diversos institutos jurídicos, alguns deles mencionados nominalmente, tais como os indícios, as presunções e os *standards* de prova. De outro lado, há uma série de temas fundamentais ao Direito que, embora não citados expressamente, estão por trás de diversas das afirmações aqui realizadas, tais como “epistemologia”, “critérios epistêmicos”, “teorias da verdade”, “livre convencimento motivado”, entre tantos outros.

Logo, goste-se ou não, o raciocínio indutivo é parte indissociável do cotidiano, possibilitando a tomada de decisões individuais orientadas por um acervo compartilhado de experiências. No Direito, como visto, a sua importância não é menor, sendo indispensável à avaliação de evidências com a finalidade de reconstruir os fatos em processos judiciais.

Nesse sentido, a qualidade do raciocínio probatório levado adiante pelo julgador é influenciada por duas variáveis. De um lado, pela observância aos critérios aplicáveis à formação e avaliação dos elementos probatórios que servem de base à premissa fática da sentença. Como visto, o julgador deverá respeitar os critérios de tamanho/grandeza, representatividade e aleatoriedade, a fim de assegurar a força da inferência. Nos casos em que o raciocínio depender de provas indiretas, o vínculo causal entre a ocorrência provada e aquela que se quer provar por meio dela deve ser devidamente justificado, de tal modo a evitar que se atribua caráter de causalidade à mera correlação. Em hipóteses em que a apuração dos fatos depender da demonstração da semelhança entre certas características de dois eventos, a força da indução analógica dependerá de um juízo a respeito da relevância da similaridade, do número de semelhanças e da natureza e grau das diferenças. De outro lado, pelo *standard* de prova aplicável ao caso: quanto mais rigoroso for o grau de suficiência exigido para que determinada hipótese probatória seja considerada provada, mais fortes devem ser a hipótese, os meios de prova e o grau de segurança inferencial do raciocínio probatório.

REFERÊNCIAS

- Ávila, H. (2018). Teoria da prova: standards de prova e os critérios de solidez da inferência probatória. *Revista de Processo*, 43(282), 113-139.
- Bobbio, N. (1993). *Teoria generale del diritto*. Giappichelli.
- Cabral, A. da Silva (1993). *Processo Administrativo Fiscal: Decreto n.º 70.235/72*. Saraiva.
- Chiassoni, P. (2007). *Tecnica dell'interpretazione giuridica*. Il Mulino.
- Damer, T. E. (2013). *Attacking Faulty Reasoning* (7. Ed). Wadsworth.
- Falcão, A. de Araújo (1959). *Introdução ao Direito Tributário: parte geral*. Edições Financeiras.
- Ferrer Beltrán, J. (2021). *Prueba sin convicción: estándares de prueba y debido proceso*. Marcial Pons.
- Gascón Abellán, M. (2010). *Los hechos en el derecho: bases argumentales de la prueba*. (3. ed.). Marcial Pons.
- Godden, D. e Walton, D. (2007). A Theory of Presumption for Everyday Argumentation. *Pragmatics and Cognition*, 15, pp. 313-346.
- Guastini, R. (2011). *Interpretare e argomentare*. Giuffrè.
- Guastini, R. (2015). Interpretación y Construcción Jurídica. *Isonomia*, 43, pp. 11-48.
- Hage, J. (1997). *Reasoning with rules. An essay on legal reasoning and its underlying logic*. Kluwer.
- Hellström, P. (2009). A Uniform Standard of Proof in EU Competition Proceedings. *European Competition Law Annual 2009: The Evaluation of Evidence and its Judicial Review in Competition Cases*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2147705
- Hurley, P. J. e Watson, L. (2018). *A Concise Introduction to Logic* (13. Ed.). Cengage Learning.
- Kelley, D. (2014). *The art of reasoning* (4. ed.). W. W. Norton & Company.
- Larenz, K. (1997). *Metodologia da ciência do Direito* (3. Ed; Trad. de José Lamego). Fundação Calouste Gulbenkian.
- Machado Segundo, H. de Brito e Machado, R. Cavalcanti Ramos (2014). Prova e verdade em questões tributárias. Em H. de Brito Machado (coord.), *A prova em questões tributárias* (p. 152-183). Malheiros.
- Mill, J. S. (2011). *A System of Logic, Ratiocinative and Inductive* (1843; volume 2). Cambridge University Press.
- Rudinow, J. e Barry, V. E. (2007). *Invitation to Critical Thinking* (6. ed.). Wadsworth.
- Schauer, F. (1991). *Playing by the rules – A philosophical examination of rule-based decision-making in law and in life*. Clarendon Press.
- Schauer, F. (2003). *Profile, probabilities and stereotypes*. Belknap.
- Schoueri, L. E. (2015). *Direito tributário* (5. ed.). Saraiva.
- Torres, R. Lobo (1988). A proibição de analogia no Direito Tributário. *Revista de informação legislativa*, 25(100), 261-268.
- Torres, R. Lobo (2006). *Normas de interpretação e integração do Direito Tributário* (4. ed.). Renovar.
- Villa, V. (2012). *Una teoria pragmaticamente orientata dell'interpretazione giuridica*. Giappichelli.
- Wróblewski, J. (1979). *Meaning and Truth in Judicial Decision*. Juridica.
- Legislação e jurisprudência
- Brasil. Lei n. 5.172, de 25 de outubro de 1966, Código Tributário Nacional. https://www.planalto.gov.br/ccivil_03/leis/l5172compilado.htm
- Brasil. Lei n. 9.430, de 27 de dezembro de 1996, dispõe sobre a legislação tributária federal, as contribuições para a seguridade social, o processo administrativo de consulta e dá outras providências. https://www.planalto.gov.br/ccivil_03/leis/l9430.htm
- Conselho Administrativo de Recursos Fiscais (Brasil), 4ª Câmara, 1ª Turma, Acórdão n. 2401005.952, de 27 de janeiro de 2019.
- Constituição da República Federativa do Brasil de 1988*. https://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm
- Supremo Tribunal Federal (Brasil), Tribunal Pleno, Ação Direta de Inconstitucionalidade n. 4.277, de 05 de maio de 2011. <https://redir.stf.jus.br/paginadorpub/paginador.jsp?docTP=AC&docID=628635>

DELIBERACIÓN, FUNCIÓN EPISTÉMICA Y ESTABLECIMIENTO DE HECHOS. BASES DE UNA ESTRUCTURA METODOLÓGICA PARA EL TRIBUNAL PENAL*

Raúl Díaz Manosalva**

RESUMEN: La función epistémica de la deliberación y su relevancia para la calidad del razonamiento probatorio, al menos en Chile, no ha sido del todo destacada frente a la atención que sí ha tenido la valoración y decisión sobre la prueba y su justificación en la sentencia definitiva. Tomando como ejemplo el Tribunal de Juicio Oral en lo Penal chileno mostraremos que, más allá de ciertas prácticas establecidas, resulta necesaria una revisión crítica sobre la deliberación de los hechos por jueces profesionales y cómo estos arriban a las razones que sintéticamente se expresan en el veredicto, cimiento de la sentencia. Pretendemos demostrar la utilidad de sentar las bases de una metodología y proponer algunas guías que permitan organizarla, en orden a mejorar el debate y argumentos en dicho espacio colectivo.

PALABRAS CLAVE: deliberación; función epistémica; prueba; hechos; veredicto.

* Este trabajo es una versión corregida y ajustada de aquella aprobada para obtener el grado de Máster en Razonamiento Probatorio de la Universitat de Girona, cuyo tutor fue el Dr. Edgar Aguilera. Agradezco especialmente su guía y, además, al Dr. Jordi Ferrer y al Dr. Diego Dei Vecchi quienes integraron el tribunal de defensa, por las sugerencias para profundizar este tema. Mención de especial gratitud para Sebastián Bravo Ibarra, por su disposición a compartir y debatir las ideas de este trabajo.

** Abogado, Máster en Razonamiento Probatorio, Universitat de Girona y Università degli Studi di Génova; Magíster en Derecho Penal, Universidad de Chile; Juez del Segundo Tribunal de Juicio Oral en lo Penal de Santiago, Chile (rdmanosalva@gmail.com).

DELIBERATION, EPISTEMIC FUNCTION AND ESTABLISHMENT OF FACTS. BASES OF A METHODOLOGICAL STRUCTURE FOR THE CRIMINAL TRIAL COURT

ABSTRACT: The epistemic function of deliberation and its relevance for the quality of probative reasoning, at least in Chile, has not been fully highlighted compared to the attention that the assessment and decision on the evidence and its justification in the sentence. Using the Chilean Oral Criminal Trial Court as a reference, we will show that, beyond certain established practices, a critical review is needed on the deliberation of the facts by professional judges and how they arrive at the reasoning synthetically expressed in the verdict. We intend demonstrate the usefulness of laying the foundations of a methodology and propose some guides that allows its organization, in order to improve debate and arguments in said collective space.

KEYWORDS: Deliberation; epistemic function; evidence; facts; verdict.

SUMARIO: INTRODUCCIÓN Y ACLARACIONES PRELIMINARES.— 1. TRIBUNALES COLEGIADOS Y ESTABLECIMIENTO DE LA PREMISA FÁCTICA. EL CASO DEL TJOP CHILENO.— 2. ERRORES Y DEFECTOS EN LA DELIBERACIÓN Y VEREDICTO.— 3. DELIBERACIÓN, MOTIVACIÓN Y PRIMER CONTROL INTERSUBJETIVO INTRA-DECISIONAL: 3.1. Justificación de la premisa fáctica y jueces profesionales; 3.2. La deliberación y su función epistémica: método colectivo para tomar decisiones y primer control intersubjetivo entre pares; 3.3. Incremento gradual de la justificación y publicidad de las razones desde la deliberación a la sentencia.— 4. PROBLEMAS DE LA DELIBERACIÓN COLECTIVA. HERRAMIENTAS PREVENTIVAS Y CORRECTIVAS CENTRADAS EN LOS AGENTES.— 5. UNA ESTRUCTURA METODOLÓGICA PARA DELIBERAR SOBRE LOS HECHOS Y JUSTIFICARLOS EN EL VEREDICTO: 5.1 ¿Cuáles son las hipótesis sobre las que debemos pronunciarnos? Determinación de las hipótesis en disputa; 5.2. ¿Hay cuestiones no controvertidas que podemos usar para establecer lo fáctico? Verificar la existencia y determinar el valor que se dará a hechos admitidos; 5.3. ¿Cuáles son las pruebas que podemos utilizar y que aportan información útil para resolver y cuáles no?, y ¿qué inferencias podemos extraer de cada una de las pruebas y cómo se complementan? Valoración de la prueba según el sistema aplicable; 5.4. ¿Qué grado de confirmación tienen las hipótesis y qué decisión debemos tomar en consecuencia? Decisión sobre los hechos aplicando el estándar probatorio; 5.5 La confección del veredicto fundado.— 6. CONCLUSIONES.— BIBLIOGRAFÍA.

INTRODUCCIÓN Y ACLARACIONES PRELIMINARES

La justificación del establecimiento de la premisa fáctica constituye una preocupación fundamental para el razonamiento probatorio. En Chile, bastante se ha dicho y escrito sobre el despliegue argumentativo exigible en la sentencia definitiva (Accatino, 2010; Araya, 2018; Cerda, 2011 y 2016; Maturana Baeza, 2014), pero no acerca del momento en que los tribunales colegiados deciden sobre la *quaestio facti* apenas concluido el juicio. En el caso del Tribunal de Juicio Oral en lo Penal chileno —en

adelante TJOP— o similares de carácter colegiado e integrados por jueces profesionales¹, la redacción de la sentencia está precedida de una deliberación entre quienes conforman la sala, a la que sigue la comunicación de una decisión en forma de veredicto de absolución o condena. Allí el tribunal resuelve sobre los hechos² luego de la rendición de pruebas sujetas a contradicción y concentra dos de los tres momentos de la actividad probatoria: *valoración y decisión* (Ferrer, 2007, p. 41)³. Esto, por regla general, en espacios excluyentes de terceros —deliberación privada o secreta— y en un tiempo limitado por ley y agenda.

Las exigencias de justificación del juicio fáctico suelen estar definidas como requisitos de la sentencia definitiva, con una carga de expresión y fundamentación elevada⁴. Sin embargo, esos argumentos probatorios deberían nacer en el seno del tribunal en un contexto deliberativo. En efecto, el texto final es la manifestación escrita de una valoración y decisión que se tomó en forma precedente a su redacción.

La deliberación debería propiciar un análisis racional de las evidencias, de modo que la decisión probatoria obedezca a una reflexión sobre las hipótesis en disputa y el grado en que cada una ha sido corroborada o refutada. Sin embargo, el contexto generalmente privado en que se despliega impide conocer la forma en que se lleva adelante esta tarea. Estimamos que resulta necesario problematizar y someter a escrutinio este momento más allá de los límites de las salas judiciales.

Si buscamos satisfacer pretensiones compatibles con una concepción racionalista de la prueba⁵, resultaría contraproducente abandonar este espacio a un abordaje

¹ Esto delimita el trabajo a tribunales colegiados conformados por jueces profesionales, excluyendo al jurado. Si bien toma como ejemplo un tribunal penal, estas reflexiones son aplicables a cualquier otro que establezca hechos y en diversas materias, con sus reglas probatorias particulares.

² Abordaremos únicamente la justificación de la premisa fáctica, aunque se resuelva junto con la normativa y, según veremos, esta condicione el debate sobre aquella. Las dificultades específicas de la prueba ameritan un tratamiento diferenciado pues, a diferencia de lo normativo, su desarrollo y prácticas no son exclusivas de la comunidad jurídica (Coloma, 2014, pp. 407-408).

³ Quien antepone a los mencionados el momento de *conformación* del conjunto de pruebas.

⁴ Incisos 2 y 3, art. 297 del Código procesal penal de Chile (en adelante CPP): «El tribunal deberá hacerse cargo en su fundamentación de toda la prueba producida, incluso de aquella que hubiere desestimado, indicando en tal caso las razones que hubiere tenido en cuenta para hacerlo». «La valoración de la prueba en la sentencia requerirá el señalamiento del o de los medios de prueba mediante los cuales se dieron por acreditados cada uno de los hechos y circunstancias que se dieron por probados. Esta fundamentación deberá permitir la reproducción del razonamiento utilizado para alcanzar las conclusiones a que llegare la sentencia».

⁵ La concepción racionalista de la prueba tiene como tesis principales (Ferrer, 2021, pp. 17-18): a) una relación teleológica entre prueba y verdad, donde el objetivo institucional del proceso es alcanzar la verdad a través de ella; b) una noción de verdad por correspondencia de los enunciados fácticos con la realidad; c) la fiabilidad de la prueba en cuanto no otorga certeza racional y puede ser fuente de errores (falsos positivos y negativos); d) un razonamiento probatorio necesariamente probabilístico, esto es, decir que un enunciado está probado es afirmar que es probablemente verdadero (a un nivel a determinar a través de reglas de suficiencia o estándares de prueba). En dicha línea, Accatino (2019) profundiza sobre las tesis centrales de esta concepción, en sus versiones angloamericana y continental o latina.

judicial carente de algunas directrices mínimas. Aunque hay exigencias normativas de justificación del veredicto producto de la deliberación, es deseable posibilitar y profundizar un efectivo control horizontal de los argumentos entre jueces.

Estimamos que la deliberación podría ser epistémicamente más valiosa o fructífera si se cuenta con cierta organización que identifique fases ineludibles de la cuestión probatoria y fáctica. En dicha línea, propondremos una estructura que pueda servir de guía o método para deliberar sobre los hechos y presentaremos pasos básicos que, como una hoja de ruta, permitan localizar argumentos, acuerdos y disensos entre jueces y faciliten su revisión retrospectiva.

Para llegar a ese punto estimamos necesario, previamente, precisar algunos alcances teóricos sobre la motivación de la premisa fáctica y de la deliberación colegiada como una forma de justificación colectiva, el control intersubjetivo de pares o *intra-decisional* y las problemáticas del intercambio grupal.

1. TRIBUNALES COLEGIADOS Y ESTABLECIMIENTO DE LA PREMISA FÁCTICA. EL CASO DEL TJOP CHILENO

El sistema procesal penal chileno y otros afines establecen un modelo en que un juez unipersonal, de garantía, atiende el control de la investigación y la preparación del juicio, mientras que un tribunal colegiado, el TJOP, es el competente para conocer del juicio oral.

Terminado un juicio, el tribunal señalado delibera en privado sobre las pruebas hasta alcanzar un veredicto de absolución o condena, paso anterior a la redacción de la sentencia definitiva. Este veredicto debe comunicarse a las partes, señalando los «fundamentos principales tomados en consideración» para llegar a la decisión, obligando a justificarla (*cf.* Arts. 339, 343 y 344 del CPP).

Esta división entre deliberación, veredicto y sentencia en el caso chileno está normativamente diseñada y parece ser la regla común en tribunales colegiados de integración profesional de otras latitudes latinoamericanas⁶, que exigen comunicar una decisión de absolución o condena que contenga sucintamente sus fundamentos.

A la hora de examinar el juicio oral, la doctrina procesal chilena (Cerda, 2003; Horvitz y López, 2004, pp. 336 a 338; Maturana y Montero, 2017, pp. 985 y 992) no aborda la deliberación sobre los hechos más allá de un acercamiento normativo formal, como la nulidad en caso de omisión de la comunicación de veredicto en el plazo legal y la formación de los acuerdos (Horvitz y López, 2004, pp. 209 y 210, mencionan los arts. 19, 72, 81, 83, 84 y 89 del Código orgánico de tribunales de Chile; Maturana y Montero, 2017, p. 993), sin que se advierta una reflexión relativa al modo en que en tal espacio se adoptan las decisiones. La privacidad de esta ins-

⁶ Por ejemplo, arts. 360 y 364 del Código procesal penal de Costa Rica y 392 y 396 Código procesal penal de Perú. En la misma línea, arts. 319 y 324 Código procesal penal modelo para Iberoamérica.

tancia⁷ y la falta de medios de revisión e impugnación del veredicto podría explicar la falta de interés en profundizar sobre esto desde un sentido práctico. En concreto, el foro procesal penal sitúa el tratamiento judicial de la prueba en la valoración y decisión en la sentencia, conforme a sus requerimientos legales.

También se debe considerar que la lectura tradicional del acercamiento judicial a la prueba y los hechos lo ocupa un juez unipersonal. Amaya (2020, p. 141) advierte sobre este enfoque individualista de los procesos de toma de decisiones. En el caso chileno esto se ve acentuado por la falta de otros tribunales colegiados de primera instancia⁸ y que instituciones como el jurado son ajenas a nuestro ordenamiento. La propia comprensión judicial de la deliberación y su producto, el veredicto, es en algunos casos confusa o derechamente errada, al concebirla como un momento en que no se deberían valorar pruebas ni justificar el establecimiento de los hechos⁹.

Ahora, en la práctica judicial se aprecia una tendencia a deliberaciones cada vez más breves. En base a información disponible sobre TJOP de la Región Metropolitana, se puede observar que el tiempo promedio de minutos formales destinado a deliberación —desde que el tribunal se retira a debatir y comunica el veredicto— para juicios de una sola jornada de duración, tuvo un descenso del 68 por 100 entre el año 2018 y el 2023. Ese mismo promedio de minutos tuvo una baja de un 31 por 100 si se miden juicios de dos o más jornadas de duración, que por volumen de prueba y complejidad requerirían mayor tiempo de análisis¹⁰.

Por cierto, los minutos señalados no cuentan como tiempo real o efectivo de deliberación y son solo un indicador de un descenso en el tiempo formalmente establecido para dicha tarea. Es efectivo que existen espacios para que los jueces deliberen paralelo al desarrollo del juicio y adopten posiciones preliminares —aun con información incompleta— a consolidar una vez finalizado. Sin embargo, podría ser que se

⁷ No es el objetivo de este trabajo abogar por la publicidad de la deliberación. Se asume su carácter secreto como un elemento que dificulta su examen crítico.

⁸ La integración plural se emplea en segunda instancia y casación y en ciertas judicaturas especiales, como los Tribunales Ambientales o de Libre Competencia, donde también se establecen hechos por composición colegiada.

⁹ Manifiesto en SCS 26152/2023, de 28 de abril de 2023. Reprocha, por otras buenas razones, parcialidad del tribunal al referirse a testigos de la defensa, pero sostiene, erradamente, que el espacio para valorar es sólo la sentencia y no el veredicto post deliberación (considerando octavo): «el principio contradictorio que rige en nuestro sistema procesal penal constituye un elemento del derecho a defensa y, asimismo, del debido proceso, por cuanto la neutralidad del tribunal se asegura y garantiza, lo que, en la especie, fue vulnerado al adelantar y emitir opinión en la etapa del pronunciamiento del veredicto, antes de la sentencia».

¹⁰ Promedio de elaboración propia para este trabajo a partir de reportes mensuales de ocho TJOP facilitados por la coordinación del Centro de Justicia de Santiago, con datos disponibles desde 2018, pero no publicados. Para juicios de una jornada de duración, el tiempo promedio bajó de cuarenta y un a trece minutos desde 2018 a 2023; en juicios de dos o más jornadas de duración, cayó de noventa y seis minutos en 2018 a sesenta y seis en 2023. Hoy es frecuente entregar veredictos verbales y su escrituración, que podría incidir en mayores tiempos, se ha reservado a casos complejos.

esté privilegiando decidir por sobre argumentar y que, sin deliberar mayormente, sea el ponente quien fundamente en forma individual la decisión ya tomada. Lo advertía Calamandrei (1960, p. 124) al hablar de hipocresía formal o de la motivación como disfraz lógico de los móviles de la decisión.

La baja podría explicarse también por el aumento de juicios y demandas de gestión, debates más eficientes producto de una mayor experiencia o derechamente porque dicho espacio esté siendo mal utilizado. Lo relevante es destacar que los jueces deliberan privadamente de una forma delineada de acuerdo con prácticas formadas en el ejercicio diario, pero desconocidas para aquellos ajenos a este. Si bien no todos los casos demandan intensas discusiones por su menor dificultad probatoria, esta afirmación debe ser matizada tanto por la existencia de casos difíciles o complejos¹¹, como por la índole extremadamente relevante de los valores en juego en la decisión, que obligan a un ejercicio judicial cuidadoso en todo caso que requiera pronunciamiento sobre la *quaestio facti*¹².

Aunque no necesariamente del indicio del descenso en el tiempo se sigue uno en la calidad de la deliberación, dado el panorama descrito, que asigna poca relevancia y no problematiza esta etapa, podemos sostener una fundada preocupación por la eventual merma en la confrontación de argumentos sobre la prueba al interior del tribunal colegiado.

2. ERRORES Y DEFECTOS EN LA DELIBERACIÓN Y VEREDICTO

Hasta aquí lo único que parece claro es que enfrentado a la deliberación el tribunal cuenta con las evidencias presentadas, un sistema de valoración y una regla de decisión o estándar probatorio (sana crítica y duda razonable, respectivamente). Salvo algunas normas sobre precedencia de votos o de formación de mayorías ante dispersión de opiniones, no se regula una estructura o pasos necesarios para organizarla.

¿Por qué debería preocuparnos la deliberación? Principalmente porque éste es el momento institucionalmente diseñado para que el tribunal colegiado valore y decida. Además, porque está establecida como un mecanismo para alcanzar decisiones sobre la veracidad de los hechos discutidos, a través de la confrontación dialogada de argumentos. Para Amaya (2020, p. 142) la deliberación sobre los hechos en el ámbito del

¹¹ La distinción de MacCormick entre *casos fáciles y difíciles* en el ámbito de la teoría de la argumentación jurídica es destacada por Amaya (2017, p. 24) por su utilidad para el desarrollo de una teoría del razonamiento probatorio en el derecho. Problemas de *interpretación, relevancia, prueba y clasificación* —o *calificación*— configurarían un caso difícil; aunque los dos últimos tratarían específicamente sobre los hechos, explica cómo los otros también podrían llegar también a afectar su determinación.

¹² El manual del curso habilitante para jueces penales de la Academia Judicial de Chile (Gajardo y Hermsilla, 2020, p. 307 y ss.) propone pautas que estimamos insuficientes. Resulta llamativa esta reflexión: «El legislador se cuidó en extremo de dar abundantes normas sobre el particular, entendiendo esta actuación de la deliberación como un trámite más o menos rápido que refleje de alguna manera el cambio del antiguo sistema...», esto último en referencia al procedimiento inquisitivo.

derecho tiene por finalidad principal la adquisición de creencias verdaderas, mediante un proceso de intercambio de razones y Ferrer (2005, p. 55 y ss. y 2007, pp. 29-32) indica que la verdad es el fin institucional y objetivo fundamental de la actividad probatoria en el proceso. Un ejercicio deficiente hará que no cumpla su función epistémica de ser una instancia racional y profunda de análisis colectivo de la prueba.

En esta fase se podrían producir diversos errores relativos a la prueba y los hechos que, independiente de si se cuenta o no con recursos para impugnarlos, se reflejarán en la sentencia al tener origen en una deliberación defectuosa, considerando que el razonamiento probatorio que le sirve de base nace aquí.

Carbonell (2022, pp. 66-73), señala que una decisión judicial entendida tanto como *actividad* o como *resultado*, puede ser defectuosa por la existencia de errores (también abusos). La autora focaliza en aquellos que pueden ser objeto de impugnación y agrupa en una tipología aquellos relacionados con *defectos de justificación externa de la premisa fáctica*: problemas generales en la determinación de tal premisa (se valora mal), construcción deficiente de las inferencias probatorias, ignorancia de algún medio probatorio, contradicciones en la valoración y error en la aplicación del estándar de prueba; estos los asocia con defectos específicos de la sentencia y causales del recurso de nulidad. Lo que queremos puntualizar es que muchos tendrán origen al deliberar.

Sin pretender ser exhaustivo ni excluyente, podemos distinguir errores y defectos *orgánicos* o *procesales* al deliberar o emitir un veredicto, de aquellos errores y defectos propiamente *epistémicos*, referidos a la adquisición de conocimiento en el espacio colectivo de la deliberación y su justificación.

Respecto de los *orgánicos* o *procesales* (no epistémicos), encontramos al deliberar defectos de participación (no todos los jueces deliberan; *v.g.*, un juez no vota o se ausenta del debate) y, al dictar el veredicto, cuando no se entrega dentro de plazo— caso sancionado expresamente con la nulidad del juicio—, es inmotivado en relación con la premisa fáctica o incompleto, al no pronunciarse sobre todas las hipótesis de hechos planteadas.

Sobre los defectos y errores propiamente *epistémicos*, que inciden en la formación de la justificación de los hechos probados o no en la deliberación y que se plasmarán en el veredicto, podríamos distinguir:

a) Errores *deliberativos*, referidos a la formación o adquisición de conocimiento en el espacio colectivo: *v.g.*, no hay confrontación de argumentos sino una mera suma de posturas individuales o interfieren sesgos grupales.

b) Errores *metodológicos* en la aplicación del marco teórico y estructural que permite el desarrollo u organización de una deliberación sobre los hechos, al ordenar, dotar de contenido u omitir pasos de la construcción de la premisa fáctica; *v.g.*, cuando no se identifican todas las hipótesis del caso o los hechos admitidos o controvertidos, no se realiza una valoración individual de la prueba antes de la conjunta o tal análisis no se separa de la decisión sobre aquella.

c) Errores en los *argumentos probatorios*, referidos en específico a las inferencias probatorias y a la valoración de los elementos de juicio propiamente tales y la aplicación de las reglas de decisión; *v.g.*, se utilizan máximas de la experiencia infundadas, se sobrevaloran informes periciales o se aplica incorrectamente el estándar de prueba.

En este trabajo nos enfocaremos en aquellos errores *metodológicos* o estructurales de la deliberación, por su trascendencia en el tratamiento de la prueba y su conexión con la argumentación en la sentencia. De esta forma se podrían evitar defectos al deliberar y también veredictos inmotivados o incompletos. Es esta preocupación la que, a nuestro juicio, amerita visitar la deliberación y visibilizar problemas específicos que pueden surgir allí, posibles correcciones apoyadas en el control entre pares y explorar guías que permitan sentar bases de un método común a emplear.

3. DELIBERACIÓN, MOTIVACIÓN Y PRIMER CONTROL INTERSUBJETIVO INTRADECISIONAL

3.1. Justificación de la premisa fáctica y jueces profesionales

Una cuestión fundamental para un ejercicio probatorio racional es la obligación de motivar el establecimiento de la premisa fáctica, entendida como «expresión lingüística de las razones que justifican la decisión adoptada» (Ferrer, 2021, p.189), en el sentido de *tener y dar razones* para decidir de un modo determinado, que actúan como garantía del respeto al debido proceso y permiten su control por las partes y revisores. Los jueces deben justificar su decisión pues de allí emana su legitimidad¹³ y precisamente esto sería, junto con una obligación, la ventaja de los jueces profesionales frente a un jurado que no motiva —al menos a un mismo nivel¹⁴— su decisión sobre los hechos en el veredicto, volviéndola inexpugnable o difícilmente controlable¹⁵.

Resulta necesario que los jueces expliciten el modo en que valoraron la prueba individual y conjuntamente y estimaron superado o no un determinado estándar de suficiencia probatoria, para vencer a la presunción de inocencia en su faceta de regla

¹³ Duce y Riego (2012, pp. 390-392) estiman consistente histórica y culturalmente con la tradición continental la opción chilena por jueces profesionales y destacan que su legitimidad está en la motivación y fundamentación de sus decisiones, a diferencia de la representación popular del jurado. Taruffo (2010, pp. 206-208) aborda las críticas a las capacidades de los jueces profesionales para enfrentar funciones epistémicas y los argumentos que la contrarrestan.

¹⁴ Por ejemplo, en España la Ley Orgánica 5/1995, de 22 de mayo dispone, en su art. 61.1 e), que el acta de votación del jurado debe contener un apartado con los elementos que formaron su convicción y una sucinta explicación de las razones para declarar o rechazar declarar hechos como probados.

¹⁵ Por esto la cuestionable convencionalidad del jurado declarada por la Corte Interamericana de Derechos Humanos en sentencia *V.R.P., V.P.C. y otros vs. Nicaragua* de 8 de marzo de 2018, como explica Ferrer (2020, p. 372 y ss.). Enfatiza que el derecho a la prueba, integrante del debido proceso, incluye una decisión motivada sobre los hechos propia de una concepción racional —y no persuasiva— de la prueba.

de juicio y así pueda tener espacio el control intersubjetivo de los criterios de racionalidad epistémica aplicados.

De lo dicho derivan las dos dimensiones que se le reconocen a la motivación: una *endoprocésal*, que alude al control de partes y revisores de los fundamentos de la decisión y una *extraprocésal*, que extiende ese escrutinio a la sociedad toda en un contexto de ejercicio de poder democrático y público (Ferrer, 2021, pp. 189-191; Igartua, 1998, p. 180; Accatino 2003, p. 35). Aunque se esté pensando en la motivación del texto final, la expresión de motivos en el veredicto previo cumpliría con ambas dimensiones, en cuanto dirigida a las partes y a la ciudadanía¹⁶ y pese a que no pueda ser revisada vía impugnación.

Sin perjuicio, porque la integración colegiada lo favorece, quisiéramos poner el acento en el esencial examen horizontal de la argumentación y de los criterios de racionalidad epistémica de la valoración y decisión sobre la prueba entre los propios jueces que integran la sala. Eso lo llamaremos, en el contexto de la deliberación, primer control intersubjetivo de pares o *intradecisional*. Esto, pues más allá de lo que determine el derecho positivo, deliberar por definición consiste en «considerar atenta y detenidamente el pro y el contra de los motivos de una decisión, antes de adoptarla, y la razón o sinrazón de los votos antes de emitirlos»¹⁷ y no en un simple concurso de voluntades.

Sostenemos que la deliberación de un tribunal colegiado es distinta a aquella introspectiva de cada juez individual, pues deberá explicitar su posición con un sentido concreto: someterla al examen del resto. Este es un primer ejercicio de motivación, justificar ante los pares. En una segunda instancia, la deliberación contempla un deber explícito de motivar pues su producto, el veredicto, exige enunciar los fundamentos que se tuvieron en vista para establecer los hechos en disputa, que se discutieron en el debate conjunto y que constituyen la justificación colegiada.

Es cierto que desde un punto de vista lógico la deliberación no es una condición necesaria de la motivación, ya que puede existir esta sin aquella y de igual modo, podría darse una deliberación profunda y completa, pero que la decisión no esté motivada. Deliberar y motivar son ejercicios conceptualmente distintos, pero ya deliberar exige justificar ante otro. Y, en el caso chileno, el tribunal está obligado a expresar esos fundamentos en el veredicto. Sin embargo, más allá de las exigencias normativas, lo que queremos destacar es que la deliberación contribuye a la motivación o justificación de la premisa fáctica al construir colectiva y racionalmente los argumentos probatorios que sustentan su establecimiento. Esa es su función epistémica, como profundizaremos.

¹⁶ Demuestra la relevancia del veredicto considerando la presencia de partes no letradas —acusado y víctima— o en casos de connotación pública.

¹⁷ Definición de la RAE.

3.2. La deliberación y su función epistémica: método colectivo para tomar decisiones y primer control intersubjetivo intradecisional entre pares

Por la gravedad de los asuntos y sanciones aparejadas¹⁸ el legislador chileno establece que algunos casos penales sean resueltos por un tribunal colegiado. La opción por este respondería a la idea de que ofrecería mejores condiciones para resolver. Sin embargo, no es la sola pluralidad de miembros —en los términos del conocido Teorema de Condorcet¹⁹— el que agrega valor epistémico a la decisión sobre lo fáctico, sino el hecho que entre ellos deban confrontar ideas. Precisamente el espacio de deliberación abre tal oportunidad de contraste de puntos de vista para la construcción de una resolución colectiva que, «aunque pueda descomponerse en decisiones individuales que formarán unanimidad o mayoría, estarán moldeadas en dicha instancia» (Coloma, 2014, pp. 416-417).

Como precisa Martí (2006, p. 82) las resoluciones de un tribunal colegiado son *decisiones colectivas públicas* —que adoptan forma de norma jurídica o que pretenden influir en la producción de una— y *de aplicación*, que se diferencian de aquellas *decisiones colectivas públicas políticas* que se toman con un cierto grado de discrecionalidad y que son objeto del modelo de *deliberación democrática*. Sin embargo, dos principios base, estructurales y constitutivos de la deliberación mencionados por el autor, son a nuestro juicio comunes a ambas y están en línea con lo expuesto hasta aquí: el de *argumentación*, en cuanto la deliberación es un proceso discursivo en el que «existe comunicación con el objetivo de transformar las preferencias de los demás» de manera razonada, a diferencia de otras formas como la negociación o la persuasión retórica; y el principio de *procedimiento colectivo*, esto es, «un proceso de reflexión dialógica en que las diversas propuestas, los diversos argumentos, y las diversas evaluaciones de cada argumento se cruzan intersubjetivamente»²⁰.

¹⁸ Delitos sancionados con penas superiores a quinientos cuarenta y un días de presidio. Desde lo procesal, trajo aparejado un régimen recursivo distinto, suprimiendo la apelación contra sentencias de TJOP e instaurando un recurso de nulidad de causales estrictas; *cfr.* Araya (2018, p. 45); Del Río (2012, p. 272).

¹⁹ «La probabilidad de que una decisión correcta sea respaldada por una mayoría de votantes aumenta en función del número de participantes así como de la competencia epistémica de los mismos, tendiendo a 1 a medida que el número de participantes tiende al infinito...». Martí (2006, pp. 185-186) sobre dicho teorema y sus condiciones.

²⁰ Martí (2006, pp. 90-92) define la *democracia deliberativa* como un modelo de toma de decisiones político normativo, cuya propuesta básica es que las decisiones políticas sean adoptadas mediante un procedimiento democrático cuyo componente deliberativo toma una forma dialógica o discursiva, de comunicación colectiva y reflexiva, basado en el intercambio de razones que cuentan como argumentos en favor o en contra de una o más propuestas determinadas, con la finalidad de convencer racionalmente a los demás, y en el que los participantes persiguen la imparcialidad en sus juicios y valoraciones (pp. 22-24). Esto permite diferenciarlo de modelos alternativos basados en el voto —democracia como mercado— o en la negociación —democracia pluralista— (pp. 66-70).

La función epistémica de la deliberación queda de manifiesto cuando entendemos que la colegiatura del tribunal es relevante por su dimensión cualitativa por sobre la cuantitativa. Podría sostenerse que un mayor número de jueces individualmente considerados que sumen sus decisiones las haga mejores, pero acá destacamos la relevancia cualitativa, ya que la confrontación de una variedad de argumentos es lo valioso: que exista tal contraste colectivo es, idealmente y en general, epistémicamente mejor a que no, al permitirnos confirmar, refutar y someter a crítica nuestras posturas individuales²¹. Por cierto, esto no descarta que en casos particulares aun deliberando tal valor no se produzca, pero ello no invalida la idea general.

Deliberar es un modo de debatir —argumentar en forma dialogada— que se distingue por su carácter *colaborativo* y *temático* en búsqueda de resolver un problema²². Exige que la actitud de los participantes sea abierta y se reconozcan capacidades intelectuales mutuas y suficientes. Es algo distinto a la mera suma de votos individuales ausente de diálogo y confrontación de posturas. Hay un plus de rendimiento que puede obtenerse en esa actividad colectiva, en cuanto todos los participantes atribuyen mayor valor epistémico al debate cooperativo que posibilitaría una mejor decisión (Aguiló, 2021, p. 22). Amaya (2020, p. 142) habla de sinergia positiva del grupo²³. La ventaja de la pluralidad de miembros es permitir que las reflexiones sobre la prueba se expliciten y pueden ser precisadas, complementadas y refutadas por el resto.

Los jueces del tribunal penal colegiado son *pares epistémicos*, pues todos observaron la misma evidencia y contarían con iguales capacidades para valorarla y decidir sobre ella²⁴. No hay una posición o estatus distinto, reflexiones u opiniones que valgan más o menos por el sujeto/juez que las emite²⁵. Solo debería atenderse a la calidad del argumento y su conexión con la prueba.

²¹ Tuzet (2021, p. 30) menciona como uno de los criterios de control propios del contexto de justificación la colegialidad del juicio, que permite confrontar y examinar las precomprensiones de cada juez singularmente considerado.

²² Aguiló (2021, pp. 18-22) distingue *colaborativo* o *cooperativo* en oposición a *conflictivo*, según su compatibilidad o no con el objetivo perseguido por los partícipes y *temático* en oposición a *actoral*, según ponga el acento en el objeto del debate y no en los sujetos que participan en él. De la combinación surgen cuatro modos de debatir: *disputa* (conflictivo y actoral); *controversia* (conflictivo y temático); *deliberación* (cooperativo y temático) y *consenso* (cooperativo y actoral).

²³ Posner (2011, pp. 333-334) controvierte esta idea en su análisis sobre decisiones de la Corte Suprema estadounidense.

²⁴ Un concepto general en Lo Guercio (2018, p. 83): dos sujetos son pares epistémicos con respecto a una cuestión p «si ninguno de los dos exhibe una ventaja epistémica sobre el otro en relación con p, ya sea porque tuviera más o mejor evidencia o porque ostentara mayores virtudes cognitivas» (igualdad *evidencial* y *cognitiva*). Coloma (2014, p. 415) estima que hay razones, más pragmáticas que epistémicas, asociadas a una práctica frecuente de resolución de casos sobre prueba de los hechos, que llevan *prima facie* a tratar como pares epistémicos a los miembros del tribunal.

²⁵ En materia pericial se utiliza también esta noción. Los expertos podrían o no considerarse entre sí o por el juez como pares epistémicos sobre un tema en particular; sin embargo, esta idea podría resultar insuficiente cuando los desacuerdos surjan de preconcepciones o predisposiciones teórico-conceptuales, como la procedencia de distintas escuelas de pensamiento (Vázquez, 2018, pp. 84-85). Trasladado a

Precisamente por esto pueden existir discrepancias internas que subsistan incluso hasta el veredicto, tomando forma de disidencia o voto particular. En tal escenario la deliberación se transformará en una controversia en el tribunal, donde lo relevante será mostrar qué separa a los interlocutores jueces y por qué una solución se estima mejor que la posición contraria (Aguiló, 2021, pp. 27-28). Esto vuelve a relevar la función endoprocésal y extraprocésal de la motivación al permitir evaluar las distintas soluciones, cuestión que importa porque no resultaría admisible, en sistemas de decisión por mayoría, un debate consensuado dirigido a mantener la unidad del tribunal al ir contra la sujeción al derecho, la imparcialidad y la independencia²⁶.

Lo dicho muestra que al deliberar se configura una dialéctica dada no sólo por las posturas rivales de las partes en el proceso²⁷, sino que también interna entre los decisores y sus argumentos. Por lo mismo, es necesario no solo que la eventual minoría exponga sus diferencias ante la mayoría, sino que también esta se haga expresamente cargo de esas críticas, para así fortalecer la argumentación probatoria y mostrar en el veredicto el debate concreto que tuvo lugar al deliberar. Si solo la disidencia cuestiona el fallo mayoritario —normalmente con baja intensidad y malentendida deferencia— pero no en sentido inverso, se omitirá parte del fundamento de la decisión y/o mostrará que esas razones fueron irrelevantes o no surgieron en la discusión. En los primeros supuestos no se honrará el debate real y en el último la minoría no habrá puesto esas razones al servicio de la deliberación colectiva. Las disidencias aportan sofisticación a los argumentos y el veredicto debería dar cuenta de estas diferencias de modo explícito.

Además, es necesario comprender que en tribunales que funcionan con regla de mayoría las disidencias operan como una alerta sobre un eventual error en un punto de vista ya sea respecto de la valoración de elementos de juicio como de la decisión final sobre lo fáctico; esto obliga a examinar nuevamente la posición, que puede ser abandonada disipando el desacuerdo o verse fortalecida con la construcción de un nuevo argumento que conteste esa crítica, pues la existencia de desacuerdos entre jueces no constituye necesariamente una razón para *suspender creencias* (Coloma, 2014, pp. 413-418).

Todo esto conforma un control intersubjetivo *intradecisional* de la justificación del establecimiento de los hechos: intersubjetivo²⁸, pues las razones del tribunal para

jueces, pudiera ser problemático si adscribiesen a una concepción distinta sobre la prueba, racionalista o persuasiva, por ejemplo, o cuando posean una formación en materias probatorias especialmente dispar o asimétrica.

²⁶ Aguiló, (2021, p. 25): «Las tesis de fondo consideradas correctas no pueden ser soslayadas en aras de fomentar la unidad del tribunal, la buena relación entre los compañeros o la imagen que el tribunal proyecta al exterior». Esto, a diferencia de la regla de unanimidad del jurado, cuya búsqueda admitiría consensos para evitar el bloqueo de la decisión (Damaska, 2015, pp. 51-54).

²⁷ Igartua (1998, pp. 190-191); Accatino (2010, p. 124), denomina expresamente su modelo de justificación de los hechos en la sentencia penal como *analítico y dialéctico*; esto último importa tratar todas las pruebas y contrapruebas rendidas en un proceso que se estructura de modo adversarial.

²⁸ Ferrater Mora (1964, pp. 984-986): intersubjetividad descrita como «puente tendido entre la pura subjetividad y la pura objetividad», entre solipsismo y objetivismo radical. En el primer caso la validez

darlos por establecidos deben considerar argumentos racionales aportados por individuos, pero compartidos y contruidos colectiva o socialmente²⁹; *intradecisional*, pues se da al interior de un espacio deliberativo en el cual cada uno de los juzgadores, en igualdad de condiciones, deben someter a revisión la argumentación del otro y la propia, con rigor recíproco.

3.3. Incremento gradual de la justificación y publicidad de las razones desde la deliberación a la sentencia

La distinción entre *contexto de descubrimiento* y *contexto de justificación* (Atienza, 2005, pp. 4-7; Gascón y García, 2017, pp. 134-144), muestra la relevancia de las razones justificativas de una decisión por sobre aquellas razones explicativas que podrían incluir preferencias de los jueces, prejuicios y motivaciones psicológicas de diversa índole, distantes de las reglas con fundamento epistémico propias de una concepción racional de la prueba. También podría presentarse la tendencia a un razonamiento intuitivo³⁰ por sobre el deliberativo, una elección rápida condicionada por los tiempos disponibles, primando posiciones preestablecidas en función de la práctica y repetición de casos que se resuelven mecánica y/o comparativamente.

Si se estima que lo relevante son únicamente las razones que consten en un texto definitivo como la sentencia, propias del contexto de justificación, se podría fomentar que en los escenarios recién descritos la decisión preceda a la motivación y que el fallo a redactar, por un solo juez, sea el espacio en que se elabore todo o gran parte de la fundamentación. Esto, en la medida que los jueces no transparenten los motivos de sus elecciones ante sus pares y/o que la celeridad derive en ausencia de revisión de argumentos. En tal sentido, usar «veredictos tipo» contruidos con cláusulas de estilo, resulta insatisfactorio y puede reflejar ausencia de deliberación, más aún cuando se contrastan con las extensas sentencias que les siguen.

Ahora, no se trata de un problema de orden cronológico entre argumentar/decidir o decidir/argumentar. Perfectamente dentro de la deliberación podrían los jueces

de las proposiciones es relativa a un solo sujeto, negando la posibilidad de conocimiento objetivo en cuanto válido para cualquier sujeto; en el segundo caso, dado que prescinde «*del punto de vista del sujeto y se defiende un objetivismo radical, no hay propiamente conocimiento —que es siempre conocimiento poseído por sujetos— sino únicamente “lo conocido”*».

²⁹ Lo explica González Lagier (2003a, p. 23): La objetividad en sentido ontológico de los hechos externos —que su existencia no dependa del observador— es insuficiente para el conocimiento, al no asegurar que nuestro conocimiento de ellos sea objetivo. Ahora, los hechos percibidos son epistemológicamente subjetivos, en cuanto relativos a una capacidad sensorial individual. De allí que, al compartir capacidades y limitaciones de percepción, «características de nuestra especie, podemos convertir ese defecto en virtud y hablar, no de subjetividad, sino de intersubjetividad».

³⁰ Bravo (2022, pp. 38-40) destaca que la neurociencia distingue el pensamiento *deliberativo*, lento, razonado y que requiere de ciertas operaciones lógicas mínimas, del *intuitivo*, rápido, espontáneo, que no requiere mayor análisis y que fluye naturalmente. En los jueces, como en cualquier persona, el segundo tiene fuerte incidencia.

comenzar manifestando una posición preliminar y luego justificarla ante sus pares argumentativamente. Lo importante a nuestro criterio es que tales argumentos probatorios sean objeto de debate interno con el fin de que la decisión que se comunique finalmente sea aquella sometida al escrutinio colectivo.

Precisamente se espera que las posturas individuales de los jueces se sometan a crítica en la deliberación y que las ideas que se exterioricen a través del lenguaje se confronten frente a los pares. En este camino puede haber avances y retrocesos, planteamientos que son desestimados y otros que se consoliden. Pero dado que a este nivel opera el *control intersubjetivo intradecisional* de los argumentos probatorios, tales posiciones intuitivas, precipitadas incluso y que puedan emitirse en ese ámbito privado, deberían dar paso a inferencias probatorias racionales que se plasmen en la justificación de la decisión que adquirirá forma de veredicto, primer paso de la construcción de la sentencia definitiva en una escala de incremento progresivo de la motivación.

A esta altura, podemos trazar una distinción —nuevamente, sin pretensión de exhaustividad ni exclusión, solo ilustrativa— según exteriorización/exigencia de la justificación, útil para describir los tres momentos en que se adopta la decisión de la premisa menor o fáctica del silogismo judicial:

—Deliberación: *privada y progresiva*, momento que supone justificar sólo entre miembros del tribunal; admitiría aportes menos profundos por su carácter tentativo y provisorio a ser depurados hasta la adopción de conclusiones parciales y la toma de decisión individual de cada juez producto del debate;

—Veredicto: *pública y sintética*, momento en que se alcanza una decisión colectiva sobre los hechos al menos por regla de mayoría, con una carga medianamente exigente de justificación a terceros, circunscrita a la comunicación abreviada de los fundamentos que cimentarán la sentencia definitiva;

—Sentencia: *pública y exhaustiva*, que obliga a la expresión completa de los fundamentos sobre el establecimiento de los hechos en un texto definitivo desarrollado sujeto a control.

4. PROBLEMAS DE LA DELIBERACIÓN COLECTIVA. HERRAMIENTAS PREVENTIVAS Y CORRECTIVAS CENTRADAS EN LOS AGENTES

Hasta aquí hemos hablado acerca de cómo la deliberación propendería a encontrar mejores soluciones de las que se obtendrían individualmente, mirando las ventajas y el enriquecimiento de la justificación de la premisa fáctica con la participación colectiva. Sin embargo, este análisis sería incompleto si no contemplara los aspectos problemáticos inherentes a este ejercicio.

Además de los sesgos propios de cualquier juzgador individual que se enfrenta al establecimiento de los hechos mediante pruebas, la deliberación colegiada presenta

los propios, comunes a las decisiones que se abordan colectivamente. Se reconocen como principales distorsiones en la deliberación (Kahneman *et al.*, 2021, pp. 92-104; Amaya, 2020, pp. 147-148) las siguientes:

a) *Amplificación* de los sesgos cognitivos individuales una vez introducidos al grupo, sea como *multiplicación* del sesgo individual que contamina al resto o como *acumulación* de aquellos de cada partícipe;

b) *Homogenización* o reducción de la variabilidad dentro del colectivo, ya sea por efecto cascada —sumarse a las opiniones emitidas previamente descartando la propia— o por el efecto del conocimiento común —información compartida que desplaza a la individual o minoritaria—;

c) *Polarización* de posiciones o ideas que se extreman al deliberar;

d) *Dominación* como tendencia hacia la visión de los integrantes más prestigiosos o socialmente aventajados, suprimiendo las de quienes no cuentan con estas características y que se traduce en la falta de participación de los miembros o en su menor credibilidad.

Desde la *epistemología de la virtud*³¹ se ha planteado que la posesión de virtudes (intelectuales, morales, argumentativas y comunicativas) permitirían a los jueces que las posean y practiquen prevenir o corregir estos problemas. Amaya (2017, p. 23) propone un concepto de virtud como capacidad o habilidad de discernir las razones relevantes que se dan en un caso concreto y proporcionar las justificaciones correspondientes. Entonces, ciertos rasgos de carácter serían determinantes para el caso de la deliberación. Para la misma autora, las virtudes deliberativas —rasgos de carácter que promueven una buena o productiva deliberación colectiva—, además: a) facilitan que los grupos obtengan resultados epistémicamente justificados; b) contribuyen a manejar el desacuerdo y alcanzar el consenso; c) generan un entorno deliberativo que favorece que se realicen los dos objetivos anteriores (Amaya, 2021, p. 147). Sin perjuicio de que actuarían en forma conjunta, nos parecen especialmente aplicables a la deliberación las virtudes *epistémicas* o *intelectuales* como la autonomía, sobriedad y valentía intelectuales, apertura de mente, perseverancia, humildad, meticulosidad e imparcialidad³²; y virtudes *argumentativas*, como la disposición a comunicarse y escuchar a otros, a modificar la propia posición en base a los argumentos o críticas que uno recibe y a cuestionar lo obvio.

³¹ En términos generales, se trata de un modelo epistemológico que se ocupa de los agentes epistémicos y las habilidades que les posibilitan acceder exitosamente al conocimiento, más que en las creencias y sus condiciones. La posesión de ciertas virtudes intelectuales permitiría arribar a creencias justificadas y verdaderas. Se distingue el *fiabilismo*, centrado en facultades cognitivas estables como la memoria, los sentidos, el razonamiento deductivo, y el *responsabilismo*, que destaca ciertos rasgos de carácter a ejercer por el agente, como la tenacidad, humildad o sobriedad, que mejorarán su desempeño epistémico (Haddad, 2017, pp. 9-13).

³² Un detalle del contenido puede encontrarse en Amaya (2020) y serían especificaciones de virtudes morales fundamentales en los jueces: justicia, imparcialidad, templanza, valentía y honestidad (Amaya, 2017, p. 29).

Otro enfoque relativo al agente epistémico (Aguilera, 2022) a nuestro criterio aplicable al caso de jueces que deliberan, plantea que contar con pruebas es una condición necesaria más no suficiente de la justificación epistémica de una creencia³³. Por ende, si el acopio de evidencias fue deficiente o su valoración, tratamiento o procesamiento fue defectuoso, la justificación de la creencia que pretenden apoyar debe ser acreedora de evaluación negativa y rechazo. Empero, estas falencias podrían ser contrarrestadas con el cultivo de virtudes intelectuales, entre las centrales³⁴: a) el *deber de reflexionar* sobre la evidencia, como el paso de un estado teórico de creencia a uno deliberativo de verdad sobre un enunciado probatorio, mediante un ejercicio crítico y analítico de la fiabilidad de los elementos de juicio, para alcanzar conclusiones sobre ellos y la eventual necesidad de recolectar nuevas pruebas³⁵; b) la virtud de *ejercer autoridad sobre nuestras creencias*, adoptando una postura deliberativa que asimila la pregunta sobre la creencia en «p» con la de si «es verdad que p», que supone ir más allá de la creencia como estado mental y preocuparse por las razones que hacen más probablemente verdadero su contenido proposicional; c) adoptar *cautelas epistémicas* que disminuyan el riesgo de arribar a creencias falsas, como el acopio diligente de pruebas y el descarte de hipótesis alternativas relevantes³⁶.

Este acercamiento al tema desde las virtudes resulta atractivo pues podrían cultivarse en el espacio institucional existente, la deliberación previa al veredicto. Además, entre jueces resultaría muy útil para detectar dónde o cuándo se está incurriendo en alguna distorsión y hacerla explícita en el debate. Requerirían, por cierto, capacitación³⁷ y ejercicio consciente y constante para convertirse en hábito.

Ahora, la aplicación de la teoría de las virtudes al ámbito jurídico y en específico al razonamiento probatorio es objeto de críticas respecto de su capacidad de constituirse en una guía que mejore la práctica judicial real de la prueba y que no pase de ser una «idealización improductiva» —la menciona Amaya (2017, p. 32)— y, la que nos parece central, que el ejercicio de ellas no garantiza resultados o conocimientos verdaderos al trabajar sobre un conjunto de elementos de juicio que no es completo

³³ En sentido *proposicional*, contar con evidencia para creer «p», y *doxástico*, que el sujeto crea efectivamente en «p» y lo haga sobre la base de esa evidencia (Aguilera, 2022, p. 83).

³⁴ Aguilera (2022, pp. 88 y ss.) formula en concreto estas tres formas de ejercicio de virtudes como contribución a un *evidencialismo robusto*.

³⁵ El TJOP carece de iniciativa probatoria, pero podría replantearse como un *acopio de argumentos probatorios* suspender la deliberación para formular a las partes preguntas que surjan y que requieran una nueva discusión sobre las evidencias. En esa línea, Coloma (2014, p. 423).

³⁶ Como veremos a propósito del paso metodológico relativo a la aplicación del estándar de prueba «más allá de toda duda razonable», este incluye en su versión más exigente descartar siempre, y no de modo contingente, hipótesis explicativas compatibles con la inocencia del acusado.

³⁷ Una propuesta de capacitación a través del estudio de modelos ejemplares de jueces y sentencias en Amaya (2018, pp. 6-8), para quien cumplen una función educativa, aportan a la formación, evolución y desarrollo de la cultura jurídica y política, permiten teorizar sobre la excelencia en la práctica judicial y sirven como estándares normativos para el resto de la judicatura.

ni fiable (Haddad, 2017, p. 48). Estas condiciones, aun aplicando tales virtudes, no impedirían la posible adquisición de creencias falsas.

No obstante, conviene destacar que estas posturas apuntan al agente epistémico, es decir, más a quienes deliberan que al ejercicio deliberativo en sí. Discutir sobre el alcance de estas teorías excedería el objetivo de nuestro trabajo, pero al menos podríamos coincidir con la opinión que, pese a las limitaciones de un acervo probatorio imperfecto, su ejercicio proveerá de un grado mayor de justificación epistémica pues los juzgadores habrán hecho todo lo posible para alcanzar una mejor respuesta que los acerque a creencias verdaderas³⁸ dentro del escenario en que se encuentren. Que no garantice el resultado no significa que no contribuyan a tomar mejores decisiones; el esfuerzo por depurar razones justificativas en un contexto idóneo disminuirá la arbitrariedad y que los jueces se limiten a la exteriorización de su voluntad o convicción individual. Por ende, mejorar el marco institucional de la deliberación y afinar el proceso de formación de la decisión resulta primordial.

Nuestra propuesta aborda el tema desde una perspectiva que tiene como objeto el proceso y el espacio de toma de decisión judicial que supone la deliberación. Considerando que las discusiones del caso concreto podrían presentar distorsiones que afecten los argumentos probatorios, es necesario ubicar fases específicas del establecimiento de los hechos en que pudieran manifestarse y que facilitará centrar la atención del grupo a la corrección o solución de ese problema. Mediante una estructura de pasos para organizar la deliberación, buscamos proporcionar un camino u orden de construcción metodológico que permita visibilizar aquellos fuertes, los eventualmente erróneos, revisar la cadena argumentativa e identificar donde están los disensos o desacuerdos que requerirán específico debate. Lo anterior no obsta a que, dentro de este método, se desplieguen tales virtudes en favor de la mejora del resultado epistémico colectivo, actuando de modo complementario al ejercicio analítico.

5. UNA ESTRUCTURA METODOLÓGICA PARA DELIBERAR SOBRE LOS HECHOS Y JUSTIFICARLOS EN EL VEREDICTO

La forma de organizar la deliberación sobre los hechos no debería ser muy distinta de la forma de construir o redactar la sentencia, pues el veredicto es la comunicación sintética de la decisión fáctica que le sirve de base y se trata de un modo judicial de razonar compartido, que obedece a las exigencias o requisitos legales.

Existen modelos de tratamiento de la prueba y justificación de los hechos en una sentencia independiente de la materia. No obstante, como revisamos, son propias de un texto final centrado en la figura individual de un juez, redactor o ponente que la elabora en un espacio y tiempo más extendido del destinado a una deliberación.

³⁸ Como creencia irreprochable (doxástica) subjetivamente justificada (Haddad, 2017, p. 54).

Ahora, no hay obstáculo para que el esquema que presentaremos pueda ser aplicado también por un juzgador individual³⁹.

Proponemos organizar el debate colegiado sobre los hechos con base en un esquema o guía inclusiva de los aspectos principales del establecimiento de la premisa fáctica justificada, emanada de una deliberación colectiva efectiva y materializada en un veredicto fundado. Esta propuesta puede ser sintetizada en cuatro fases o etapas estructurales donde será necesario que los jueces respondan, separadamente, las siguientes preguntas:

1. ¿Cuáles son las hipótesis sobre las que debemos pronunciarnos?⁴⁰
2. ¿Hay cuestiones no controvertidas que podemos usar para establecer lo fáctico?
3. ¿Cuáles son las pruebas que podemos utilizar y que aportan información útil para resolver y cuáles no?, y ¿qué inferencias podemos extraer de cada una de las pruebas y cómo se complementan?
4. ¿Qué grado de confirmación tienen las hipótesis y qué decisión debemos tomar en consecuencia?

Esta sería una estructura analítica general, comprensiva y eficiente para el contexto de tiempo y espacio de la deliberación, útil para mantener un orden u organización metodológica sin incurrir en saltos argumentativos, construyendo paso a paso la decisión sobre los hechos y también para revisar, retrospectivamente, si la misma se encuentra justificada. A continuación, abordaremos el contenido esencial de cada una. Por cierto, cada fase supone una multiplicidad de cuestiones específicas sobre la prueba y los hechos que no podremos profundizar aquí, pero que obligarán a los jueces a estudiarlas y discutir las según las dificultades que cada caso les ofrezca.

5.1. ¿Cuáles son las hipótesis sobre las que debemos pronunciarnos? Determinación de las hipótesis en disputa

Los jueces al deliberar deben resolver peticiones de las partes sobre hechos concretos que pueden ser subsumidos en aquellos genéricos contenidos en las normas, con la pretensión de que se apliquen las consecuencias jurídicas allí previstas.

³⁹ En Bravo *et al.* (2023) se formula una propuesta desde una concepción racionalista de la prueba para establecer hechos en decisiones judiciales genéricas, a través de siete pasos metodológicos: (1) Describir la hipótesis fáctica del sujeto activo de la acción; (2) Describir la hipótesis fáctica del sujeto pasivo de la acción; (3) Despejar aquellos hechos pacíficos y/o convenciones probatorias; (4) Seleccionar aquellos elementos de juicio que permitan confirmar o refutar los aspectos centrales de las hipótesis; (5) Valorar la prueba individualmente y en su conjunto; (6) Aplicar reglas de decisión, aplicando un estándar de prueba al caso y las reglas sobre carga de la prueba; (7) Hacerse cargo de la eventual prueba irrelevante, sobreabundante o inconducente. Otro completo modelo para sentencias para tribunales penales chilenos en Accatino (2010).

⁴⁰ No hablamos solo de hipótesis controvertidas, acusatorias y defensivas expresas. Se podrían imaginar hipótesis alternativas no invocadas, según revisaremos.

En el juicio oral, la hipótesis acusatoria tomará la forma de un relato descriptivo de un hecho delictual y que contendrá circunstancias espacio temporales, conductas del sujeto activo, modalidades y eventualmente resultados. Las hipótesis de la defensa, por su parte, se encontrarán generalmente en los alegatos de apertura y clausura del juicio, pues no está obligada a plantear una teoría del caso previa al juicio e incluso puede optar por el silencio frente a la imputación (cfr: arts. 263 c) y 268 CPP).

No existe para el proceso penal chileno una resolución que predetermine cuáles son los hechos relevantes, pertinentes y controvertidos a probar, como ocurre en el proceso civil posterior a la demanda y contestación. Acá el acusador hará una estimación sobre qué aspectos intentará acreditar para satisfacer su pretensión. Sin embargo, ya que puede ser excesivamente descriptiva, imprecisa o genérica el tribunal deberá determinar, como primera fase de su análisis, cuál es el *thema probandum* u objeto específico del juicio al que se enfrenta⁴¹ de forma tal que la valoración posterior esté vinculada por dicha definición.

Un modo de examinar la prueba está delineado por la descomposición del hecho a probar en función de la premisa normativa. El debate sobre los elementos que conforman los tipos penales aplicables permite que los relatos sobre los hechos formulados puedan sintetizarse marginando cuestiones irrelevantes para la subsunción. Esto operaría como una especie de reducción a lo esencial que acotaría la deliberación en la medida que exista acuerdo sobre los presupuestos centrales para configurarlo.

Aquí se manifiesta con claridad cómo se entrelaza la cuestión fáctica con la normativa, lo que ha puesto de relieve la hermenéutica jurídica al señalar, a propósito del denominado círculo hermeneútico, que «el caso concreto guía la individuación de las normas relevantes, pero las normas, a su vez, guían la selección de los aspectos fácticos relevantes» (Tuzet, 2021, p. 37).

Cabe puntualizar que por mucho que la dogmática penal sistematice la teoría del delito y los requerimientos de cada ilicitud, no será siempre unánime, lo que puede complejizar la definición del *thema probandum*, la interpretación de los tipos y generar divergencias entre los jueces que deliberan.

Ahora, no todo es abstracción, ya que se deberán identificar los elementos aludidos dentro del relato del modo concreto propuesto. Por ejemplo, la fuerza en las cosas consistente en la fractura de una específica ventana de un inmueble, en el robo; la conducta homicida, mediante la concreta acción de propinar una puñalada en el pecho de la víctima; el engaño de la estafa, mediante el ardid consistente en aparentar solvencia y competencia en un negocio determinado; el exceso de velocidad fundante de la negligencia en el tráfico rodado.

⁴¹ Bravo (2023, pp. 64-69) distingue tres aproximaciones al concepto de *thema probandum*, aristas que en conjunto delimitan el concepto: desde una perspectiva jurídica, como cumplimiento de requisitos o presupuestos para la aplicación de una norma; una epistemológica, como hipótesis a ser confirmadas por elementos de prueba, y una lingüística, como el contenido de enunciados o proposiciones de un relato. Acá nos referimos indistintamente a todas ellas.

Siguiendo a Anderson *et al.* (2015, p. 96) la hipótesis acusatoria será el *probandum final* o principal que admite separarse en una serie de proposiciones más simples, *probanda penúltimas*, diferenciables en la formulación legal y que deben acreditarse para sustentar aquél. Existirán también *probanda intermedias*, proposiciones a acreditar que se enlazan argumentativamente en dirección a un *probandum penúltimo*. De este modo, los argumentos probatorios se estructuran como una cadena de razonamientos que irán desde los datos o evidencias a valorar hasta la hipótesis. En cada paso habrá una fuente de posibles dudas y debate colegiado.

Lo que decimos es también aplicable a las hipótesis defensivas que plantean la ocurrencia de determinados hechos en todo o parte distintos a los imputados, no punibles, justificados o sancionables de una forma morigerada. Pero también es muy importante destacar que, existiendo o no tesis defensivas concretas, pueden imaginarse otras hipótesis explicativas plausibles de los hechos que surgirán lógicamente frente a la imputación de cargo y que derivarían en la inocencia del acusado o en una situación más beneficiosa para él. Quienes deliberan podrán formular escenarios alternativos que corran paralelamente en la valoración probatoria, a fin de verificar su corroboración o refutación.

Existen modelos argumentativos que podrían ayudar en la tarea de visualizar la argumentación probatoria. Este espacio no permite detenerse en ellos profundamente, sin perjuicio de que hay uno que se aviene al desglose de las hipótesis en proposiciones, de modo analítico: el método de Wigmore. Mediante este es posible diagramar los argumentos probatorios de manera de comprobar si hay o no una cadena inferencial sólida que apoye el *probandum final*, en ascenso desde un nivel micro de cada proposición intermedia a un nivel macro (*probanda penúltimas y final*)⁴².

Independiente del modelo que se adopte, lo relevante es utilizar alguna herramienta que contribuya, como ésta, a analizar el apoyo probatorio de cada proposición o enunciado y las relaciones entre ellos. Pretendemos destacar aquí que la definición sobre las hipótesis a probar y su desglose determinará el curso posterior del análisis y puede ser la primera fuente de diferencias entre quienes deliberan. Por ello resulta recomendable comenzar por su discusión, de modo de contar con una base para las fases sucesivas y explicitarlas en el veredicto.

5.2. ¿Hay cuestiones no controvertidas que podemos usar para establecer lo fáctico? Verificar la existencia y determinar el valor que se dará a hechos admitidos

El concreto juicio puede ofrecer un panorama en que todo o parte de los hechos planteados en la acusación se encuentren admitidos o reconocidos expresamente por

⁴² Sobre el método señalado, su simplificación y ajuste en Anderson *et al.* (2015, pp. 165-188); también Gama (2011) distingue la propuesta de Wigmore tanto para el estudio como para el análisis de las pruebas, que aquí mencionamos.

el acusado o no ofrezcan controversia en la teoría del caso defensiva. Este tema engarza con el concepto de verdad al que se adhiera y la deliberación debiese contemplar este ítem como una cuestión a debatir por el tribunal.

Desde la conceptualización de la *verdad como correspondencia*, la veracidad de los enunciados sobre los hechos no puede ser sustituida por la voluntad de los intervinientes mediante acuerdos procesales o convenciones (*verdad como consenso*); será la prueba rendida la que respalde tal conclusión. Si bien no lo hará de forma categórica o irrefutable, pues no se puede descartar un posible error por falta de correspondencia con la realidad, sí lo hará con base probabilística a través de inferencias inductivas. Por esto, decir que un hecho está probado no equivale a decir que es verdadero, pese a las razones justificativas entregadas. Pero tampoco que un hecho sea pacífico o consensuado lo hace verdadero, cuando siquiera se entregan justificaciones probatorias⁴³. Si bien adherimos aquí a la primera tesis sobre la verdad, habrá que considerar que la existencia de un acuerdo sobre un hecho tampoco carece de valor: la afirmación coincidente sobre él constituye también una razón probatoria a evaluar.

En el caso chileno se permite arribar a convenciones probatorias en la preparación de juicio, esto es, que se den por acreditados ciertos hechos con el efecto de que ellos «no podrán ser discutidos en el juicio oral» (*cf.* art. 275 CPP). Otro tanto puede ocurrir en el juicio cuando declara el acusado admitiendo hechos o la defensa así lo formula en sus alegatos. Serán dos casos en que la admisión o acuerdo sobre lo fáctico formará parte de la materia a deliberar.

Es importante destacar, más en el caso penal, que los acuerdos sobre la *quaestio facti* vía convenciones o admisiones posibilita riesgos epistémicos de error. Por ello resulta crucial que tales mecanismos consensuales se regulen y apliquen en cada procedimiento equilibrando finalidades de eficiencia y facilitación del razonamiento probatorio en juicio sin alejarse de la averiguación de la verdad a través de la prueba⁴⁴.

Entonces, los jueces deberán revisar si hay hechos admitidos o no controvertidos definiendo claramente su contenido, para luego determinar cuál es el alcance o utilidad que puede dárseles en la justificación factual, dispensando o morigerando ya sea la carga del acusador y/o la argumentación probatoria del tribunal. Aparte de la cuestión conceptual, las limitaciones normativas no pueden obviarse pues la legislación procesal penal chilena impide «condenar a una persona con el solo mérito de su propia

⁴³ Taruffo (2010, pp. 141-145) rechaza la construcción de una verdad en base a la conducta de las partes (*hecho pacífico* o no controvertido, relevado de prueba y decisión judicial) pues las alegaciones sobre hechos son enunciados descriptivos con pretensión de verdad, que tendrán efecto jurídico solo si las pruebas las confirman y corresponden a acontecimientos del mundo real.

⁴⁴ Veleda (2021, pp. 261, 262 y 290) concluye que los mecanismos consensuales podrían perjudicar tal averiguación, pero serán las condiciones de implementación de cada procedimiento las que reflejarán el mayor o menor peso de esa finalidad frente a otros valores, como la eficiencia. En Bravo, *et al.* (2023, pp. 8-9) sostenemos que el juez deberá considerar los límites epistémicos y jurídicos de un hecho no controvertido o admitido y de una determinada convención, verificando si cumple con mínimos de racionalidad epistémica y jurídica que permitan usarlos como si se tratase de hechos probados y como base de inferencias probatorias para otros enunciados.

declaración»⁴⁵. Además, la presunción de inocencia como regla de juicio impone al acusador la carga de acreditación de los hechos que imputa. La discusión del tribunal colegiado en este punto será cuál, de qué calidad y sobre qué tópicos debiera versar el aporte probatorio independiente que complementará eventuales admisiones.

5.3. ¿Cuáles son las pruebas que podemos utilizar y que aportan información útil para resolver y cuáles no?, y ¿qué inferencias podemos extraer de cada una de las pruebas y cómo se complementan? Valoración de la prueba según el sistema aplicable

En esta fase debe conectarse la prueba y los hechos a través de argumentos probatorios. Está regida por el sistema de valoración de prueba que normativamente oblique al tribunal, para el caso del TJOP la *sana crítica racional* (cfr. inciso 1° art. 297 CPP). Araya (2018, pp. 148 y ss.) describe su comprensión en Chile y enfatiza que es un sistema libre, pero que debe someterse al tamiz de racionalidad dado por los parámetros de la lógica, las máximas de la experiencia y los conocimientos científicos. No contempla reglas que predeterminen el valor de los elementos de juicio —a diferencia de la prueba tasada— pero está sujeto a límites lógicos, epistémicos y empíricos que sirven también de fundamento a las inferencias (Bravo *et al.*, 2023, p. 13).

Esta etapa permite a su vez la división en otras tres subfases, que facilitan organizar el tratamiento de las pruebas y la identificación de fuentes de disenso deliberativo.

La primera consistirá en una *selección* de pruebas relevantes que por vía inferencial aporten información separadamente o en relación con otras, corroborando o refutando una hipótesis o los *probanda* en que se haya determinado dividirla. El examen de relevancia no será aquél lógico y abstracto propio de la etapa de admisibilidad probatoria, sino uno de carácter concreto de la información que proporciona cada medio de prueba⁴⁶. Existirán casos en que podrán ser desestimadas algunas evidencias que sortearon tal filtro pero que rendidas demostraron su inutilidad, permitiendo su rápido descarte por falta de conexión con la hipótesis.

También podrá discutirse respecto a la desestimación de pruebas consideradas ilícitas⁴⁷, planteamiento bastante común en la práctica local y que supondría, de

⁴⁵ Cfr. inciso 3° art. 340 CPP. La facilitación del razonamiento probatorio es promovida por el legislador concediendo una atenuante de colaboración al esclarecimiento de los hechos; cfr. art. 11 Núm. 9 del Código penal chileno.

⁴⁶ Como concretización del principio de relevancia, según el cual todas las pruebas potencialmente útiles para la determinación de los hechos debieran ser admitidas en juicio (Taruffo, 2010, p. 163). Una prueba es relevante si aporta apoyo o refutación de alguna de las hipótesis fácticas del caso a la luz de los principios generales de la lógica y la ciencia (Ferrer, 2007, p. 42).

⁴⁷ En Chile se denomina a este ejercicio de desestimación de prueba ilícita por el TJOP, que no fue excluida en la etapa intermedia, con el confuso rótulo de «valoración negativa de la prueba». Reflexiona sobre esta problemática Aristegui (2020).

accederse, marginar del conjunto de elementos rendidos en juicio por fines extrínsecos⁴⁸ aquellas que se estiman vulneradoras de garantías fundamentales y derechos constitucionales.

A continuación, conforme expone González Lagier (2003b, pp. 35-36) —aplicando al razonamiento probatorio el esquema de la Teoría de la Argumentación de Stephen Toulmin—, la valoración de la prueba propiamente tal implica que los jueces construyan *inferencias probatorias* que desde las *premisas* o medios de prueba apoyen determinados enunciados probatorios que serán las *razones* del argumento; estos a su vez se conectarán mediante un *enlace* o *garantía* expresivo de una regularidad que autorice a pasar de las *razones* a la *pretensión* o *conclusión* de la inferencia, que será la hipótesis que se pretende confirmar o refutar. Este *enlace* encontrará a su vez *respaldo*: empírico en el caso de generalizaciones como las máximas de la experiencia o los conocimientos científicamente afianzados, normativo en el caso de las presunciones y doctrinario y/o jurisprudencial para un concepto o teoría.

Para llegar a estas inferencias, el segundo paso al deliberar será la *valoración individual* de las pruebas. Este ejercicio puede superponerse con el primero, pues seleccionar implica valorar. Sin perjuicio, nos referirnos acá a la estimación del concreto apoyo que por vía de corroboración otorgue el medio de prueba seleccionado como relevante a uno o más «*probanda*» en que se descomponga la hipótesis, así como las eventuales refutaciones que pueda mostrar una contraprueba.

Al efecto, existen una serie de criterios de valoración individual, por lo general categorizados como fiabilidad, credibilidad, integridad, entre otros. Las tipologías fundamentales de medios probatorios —testimonial, pericial, documental, declaraciones de parte— permiten la agrupación de elementos de juicio considerando características comunes, pero cada una de ellas admite otras especificaciones con sus propias particularidades y problemáticas, considerando la multiplicidad de formas en que puede encontrarse información relevante para establecer lo fáctico. Incluso, la irrupción de nuevas tecnologías o soportes de datos utilizables en el ámbito probatorio podrían demandar de los jueces conocimientos específicos y originar múltiples discusiones sobre ellos.

En tercer lugar, la *valoración conjunta* de la prueba importa la consideración del apoyo recíproco que se brindan las pruebas unas con otras o las contradicciones que puedan surgir desde una perspectiva más amplia o global de las relaciones entre los medios de prueba. Los criterios aquí son distintos, al tratarse de atributos de un cúmulo de datos en relación con la hipótesis que pretenden confirmar. A modo ejemplar, González Lagier (2022a, pp. 383-394) distingue bajo la denominación de *criterios de racionalidad epistemológica*, los siguientes:

⁴⁸ En el sentido que protegen valores distintos a la determinación de la verdad de los hechos, finalidad de la prueba en el proceso. Taruffo (2008, pp. 38-42) distingue de la *relevancia* como estándar lógico de conexión entre los medios de prueba y los hechos del litigio, la *admisibilidad*, que atiende a normas jurídicas o reglas de exclusión que impiden su admisión en función de la protección de otros valores o intereses y que excepcionan al principio de relevancia.

— Criterios en relación con los hechos probatorios: a mayor cantidad, variedad, fiabilidad y pertinencia o relevancia de los elementos de prueba, mejor conformación de la hipótesis;

— Criterios sobre la conexión o enlace de la inferencia probatoria: a mayor grado de fundamentación de la generalización empírica y mayor probabilidad, frecuencia o correlación de esta, mayor confirmación y solidez de la hipótesis;

— Criterios sobre la hipótesis: la hipótesis estará más confirmada si no ha sido objeto de refutación directa ni indirecta, se confirman las hipótesis derivadas, es coherente, tiene capacidad explicativa de la mayor parte de los elementos de juicio, es simple y se eliminaron hipótesis alternativas.

Como puede advertirse, los acuerdos y divergencias deliberativas pueden ser múltiples y surgir en cada eslabón de las inferencias, esto es, respecto de los medios de prueba que les sirven de base y el enunciado extraído, el enlace que le sirve de conexión con la pretensión, su respaldo y la conclusión misma a la que se arriba.

Por último, enfatizamos que las inferencias serán las razones justificativas del establecimiento de los hechos, construidas a través de criterios racionales e intersubjetivamente controlables. Por tanto, quienes deliberan deberán discernir cuándo se trata efectivamente de argumentos, descartando opiniones que sólo simulan serlo, carentes de razones que las sostengan⁴⁹.

5.4. ¿Qué grado de confirmación tienen las hipótesis y qué decisión debemos tomar en consecuencia?

Decisión sobre los hechos aplicando el estándar probatorio

La última etapa o el cierre de una deliberación organizada sobre los hechos consistirá en adoptar una decisión sobre la prueba, en particular, identificando y aplicando el estándar de prueba correspondiente, entendido como una regla que permite decidir si una determinada hipótesis puede o no aceptarse como probada, con el grado de confirmación de que se disponga y en un determinado contexto procesal⁵⁰.

Respecto del estándar de prueba, se ha puesto de manifiesto la necesidad de contar con uno que, en forma posterior a la valoración del grado de apoyo que otorgan

⁴⁹ González Lagier (2022a, pp. 359-360) estima que es un error o argucia frecuente presentar opiniones como si fueran argumentos que, al no estar respaldadas por razones, no deben aceptarse en una deliberación racional.

⁵⁰ En Bravo *et al.* (2023, pp. 19-20) se agrega la *carga de la prueba*, entendida como regla final o de clausura que determina qué parte pierde si no hay prueba suficiente para respaldar la hipótesis fáctica que funda la aplicación de una determinada norma. En el caso penal, si el acusador no prueba su hipótesis, pierde el caso y se absolverá al no ser vencida la presunción de inocencia como regla de juicio.

las pruebas a una determinada hipótesis, permita decidir sobre la suficiencia de tal grado de apoyo para darla por probada⁵¹.

Sabemos que el debate sobre la conceptualización y contenido del estándar de prueba es una cuestión de intensa discusión en doctrina y, en consecuencia, una materia sobre la cual los jueces también asumen distintas posturas que no siempre presentan una identificación clara en la lectura de sus fallos. Este resulta un tema ineludible en la deliberación.

Una formulación objetiva de estándar de prueba debe cumplir la función de establecer un umbral de suficiencia probatoria, evitando la subjetividad de expresiones que aluden a la creencia o convicción del juez y reduciendo al máximo la imprecisión propia del lenguaje (Ferrer, 2021).

Resulta claro que la fórmula del estándar local —predominante en las legislaciones penales— que exige para condenar que «el juez que haya alcanzado la convicción, más allá de toda duda razonable», de la existencia de un delito y de la participación del acusado en él (*cf.*, art. 340 CPP), está lejana a una propuesta objetiva si no cuenta con una interpretación que determine su contenido. Ahora, se ha entendido como una exigencia especialmente elevada por los valores implicados en materia penal y por lo mismo requeriría que la prueba proporcione un alto grado de corroboración de la hipótesis de cargo y permita, a su vez, el descarte de las hipótesis que muestren al acusado como inocente, alejándose de la noción de convicción⁵².

Sin embargo, Ferrer (2021, p. 35) reconoce y advierte que «la formulación del estándar de prueba tendrá siempre un residuo de vaguedad ineliminable, propia del lenguaje ordinario», por lo que sus requisitos deben ser matizados exigiéndole que minimice tal vaguedad; así, un estándar de prueba será mejor si en su formulación «el espacio de los casos claros de cumplimiento o incumplimiento del mismo es mayor y el margen de vaguedad en el que se moverán los casos dudosos es menor».

Sin perjuicio de lo anterior, hay posturas que ven en la pretensión de objetividad una meta inalcanzable. Al respecto, González Lagier (2020, pp. 90-95) plantea que si la subjetividad que se opone a lo objetivo se le considera como arbitrariedad del juzgador —en el sentido de no aportar razones— estaríamos ante una pretensión alcanzable; pero si se opone a vaguedad, en un sentido *intensional* y *gradual*, no podrá ser reducida a efectos de frenar la discrecionalidad del juez⁵³.

⁵¹ Ferrer (2021, pp. 22-24). Además, plantea que las razones de reglas de decisión como los estándares de prueba no son epistemológicas, sino que políticas, pues también cumplen la función de distribuir los riesgos de error entre las partes; en profundidad (pp. 115-138).

⁵² Accatino (2011, p. 502) para el caso chileno ha sostenido la necesidad de contar con estándares de prueba objetivos y distanciarse de una concepción subjetiva de la decisión probatoria basada en la persuasión que la prueba genera en el juez o la convicción con que este la alcanza.

⁵³ El mismo autor (2022b, pp. 102-104) sostiene que el *fundherentismo*, tesis de justificación epistémica de Susan Haack, descartaría la posibilidad de formular estándares que reduzcan la discrecionalidad mediante una escala numérica, considerando el carácter multidimensional de la prueba y los

También Dei Vecchi (2020, pp. 32-33) se ha referido a los estándares de prueba resaltando la imposibilidad de formulación precisa y objetiva, pero agrega que la justificación de la proposición de que un hecho está probado responde no solo argumentos epistémicos, sino que también pragmáticos. Explica que «la suficiencia de las razones epistémicas en favor de las proposiciones varía según el contexto de decisión» y que «lo que determina la suficiencia o completitud de las razones epistémicas en favor de una proposición fáctica son las razones prácticas vinculadas con la acción a llevar a cabo a partir de la aceptación de esa proposición». Existirían entonces consideraciones pragmáticas —que a diferencia de las teóricas determinan qué es lo que se debe hacer— que inciden en la justificación epistémica del contenido proposicional de una creencia y en cuándo puede entenderse que la prueba es suficiente para tomar una decisión en un contexto determinado.

Ahora, con este brevísimo panorama, en el contexto de la deliberación y con las diferencias que se advierten en la dogmática, ¿cómo pareciera que están resolviendo los jueces la suficiencia de la prueba? La pregunta es relevante pues se espera que la deliberación implique que los jueces se posicionen al respecto. Hay una cuestión en la que puede coincidirse y es que, al menos en Chile, atendida la exigente carga de fundamentación que imponen los códigos procesales y por influjo de la concepción racional de la prueba, la discusión sobre criterios de valoración y decisión fáctica se ha profundizado, obligando a los jueces a su estudio y fundamentación en los fallos.

Más allá de menciones o cláusulas de estilo en las sentencias, tanto a la hora de condenar en casos sólidos y de absolver ante insuficiencias probatorias patentes detectadas, no suele advertirse una diferenciación entre la valoración de la prueba y la aplicación del estándar. Son los casos de mayor complejidad los que demandan una justificación mucho más sofisticada, que obliga a distinguirlas en la sentencia. Aunque carecemos de un estudio empírico, esta aproximación primaria podría describir, en concreto, cómo se está abordando la decisión sobre suficiencia probatoria al deliberar: unos, en que el estándar se menciona como cita de estilo final, post valoración, ante casos claros de condena o absolución versus otros, en que la dificultad probatoria empuja a ejercicios argumentativos más intensos, definiendo qué se entiende por duda razonable y cuáles son las hipótesis alternativas. No afirmamos que esta forma de aplicar o posicionarse frente al estándar sea satisfactoria, pero sí que obedece a una práctica de los jueces. Tampoco puede desatenderse la todavía arraigada influencia en Chile de una visión persuasiva apoyada en la expresión *convicción*.

Por ejemplo, aquellos casos graves y que naturalmente generan poca prueba, ofrecerán dificultades al deliberar y podrían incluir consideraciones pragmáticas sobre el

distintos grados en que cumple con los criterios de valoración (apoyo, fiabilidad independiente y amplitud). Precisa que, si bien los defensores de los estándares no exigen tal escala, tampoco esta tesis avalaría satisfacer pretensiones más débiles, estándares determinados *a priori* con carácter general. A su juicio, sólo estamos en condiciones de utilizar criterios de justificación o valoración y exigir que el juez los use en la motivación, descartando creencias arbitrarias o carentes de justificación —que no es equivalente a subjetivas— y del resto, aquellas razonablemente bien justificadas, decir que lo están en base a expresiones graduales, siempre con un espacio inevitable de imprecisión y ambigüedad.

contexto de la decisión, es decir, sobre los reales escenarios probatorios y atender a las consecuencias, lo que está en juego; esto porque las exigencias de prueba no pueden separarse de las reales posibilidades disponibles y aunque ella sea escasa, no implicará necesariamente incumplir el estándar. Al deliberar puede influir que ciertos casos de prueba escasa dejen ilicitudes sin sanción (impunidad). Este es un tipo de error, falsa absolucón; no obstante, los falsos positivos (condenas a inocentes) es un riesgo de error que se estima más grave y que quiere distribuirse a través de un estándar exigente como el de la duda razonable; el equilibrio es frágil y creemos que los jueces, al decidir, son sensibles a estas consideraciones.

Avanzando en guías concretas para la deliberación, cabe destacar que Ferrer (2021, p. 209-225) propone tres posibles estándares con contenido conciliable con el de duda razonable. En todos ellos se exige un acervo probatorio de calidad que corrobore la hipótesis y su resistencia a refutaciones, pero se diferencian en las hipótesis alternativas que deben ser eliminadas, siempre exigentes, ordenadas en grado decreciente: 1) todas aquellas plausibles y compatibles con la inocencia del acusado o más beneficiosas para él, descartadas las meras hipótesis *ad hoc*; 2) las mismas anteriores en la medida que hayan sido alegadas por la defensa contraria; 3) idénticas hipótesis alegadas por la defensa, siempre que se haya aportado alguna prueba que le otorgue algún grado de confirmación.

La deliberación deberá abordar qué exigencia se hará al descarte de las hipótesis alternativas. Fundadamente algunos jueces podrían sostener una u otra formulación, haciéndose cargo únicamente de aquellas levantadas o alegadas por la defensa, o requiriendo el apoyo de pruebas que la sustenten⁵⁴.

El estándar más exigente, que supone descartar todas aquellas plausibles y compatibles con la inocencia del acusado, excluidas las *ad hoc*, imposibles de confirmar o refutar (Ferrer, 2021, p. 223), nos lleva a preguntarnos hasta dónde deben los jueces imaginar tales escenarios explicativos. Si el tribunal se posiciona aquí, debiera atender especialmente en su deliberación a la consideración de estas hipótesis alternativas de los hechos planteándose posibilidades concretas no alegadas, factibles de confirmar o refutar⁵⁵.

Sería útil explicitar en el veredicto cuáles fueron esos concretos escenarios que el tribunal discutió y por qué los desestimó o acogió para justificar el cumplimiento o no del estándar «más allá de toda duda razonable», demostrando la exhaustividad con que se deliberó. Lo que pudiera chocar con una idea de ejercicio colectivo y ser un

⁵⁴ El estándar uno, más exigente, parecería estar presente en el discurso de la práctica procesal penal chilena. Aunque establecer cuál de las propuestas de estándares es adecuada para cierto contexto es una decisión político-normativa (o político-valorativa) relacionada con la distribución del riesgo de error entre las partes en distintos tipos de casos, no siempre queda claro en las sentencias si hubo un ejercicio de descarte de hipótesis más allá de las discutidas en juicio.

⁵⁵ Aguilera (2022, p. 93) lo propone como una cautela epistémica, en la medida que esas hipótesis alternativas sean contextualmente relevantes, imaginando escenarios alternativos cercanos.

fraude argumentativo, sería que en la sentencia el ponente o redactor imaginara múltiples explicaciones que no fueron objeto de discusión u otras lejanas, extravagantes o descontextualizadas.

5.5. La confección del veredicto fundado

De seguir los pasos o fases anteriores, la elaboración de un veredicto fundado tendrá una estructura que permitirá registrar los fundamentos principales de la decisión sobre la cuestión factual, plasmando la justificación externa de la premisa menor del razonamiento judicial.

El veredicto deberá dar cuenta de los acuerdos y desacuerdos alcanzados, lo que implica poner el esfuerzo y foco en el momento de toma de decisión colegiada, para así vincular al ponente y facilitar una labor de redacción que refleje el trabajo colectivo. Si el veredicto adopta la estructura propuesta, la sentencia definitiva vendrá a ser un desarrollo y profundización de los mismos argumentos probatorios ya enunciados, ajustándose a la discusión y decisión tomada en la deliberación.

Una buena práctica es la escrituración del veredicto, particularmente en casos complejos, en que ya sea por la materia, cantidad de pruebas, trascendencia u otras dificultades, aconsejen entregar una resolución que se haga cargo con profundidad de la discusión sometida a su conocimiento. Esto no descarta emitir veredictos verbales en casos menos complejos, donde jueces con entrenamiento pueden entregar argumentos probatorios estructurados, de modo eficiente y sujeto a registro.

De lo que se trata es de redactar o expresar una motivación en el veredicto y no de usar meras formalidades, citando cláusulas formales sobre las normas aplicables a la valoración y decisión sobre la prueba, sin contenido argumentativo.

Como ya se ha dicho, si la obligación y meta es fundar un veredicto, ello tendrá diversas utilidades y efectos positivos: (a) obligará a que la deliberación sea profunda, una verdadera confrontación de argumentos que pueda, conscientes de eventuales distorsiones, mejorarlos a través del aporte colegiado; (b) limitará los argumentos de la sentencia a lo realmente deliberado, al honrar dicho debate y realzar su carácter fundamental y constitutivo de la decisión a plasmar en el fallo; y (c) permitirá a las partes conocer de inmediato las razones del tribunal y que los propios pares puedan controlar los fundamentos a desarrollar en la sentencia.

6. CONCLUSIONES

La deliberación del TJOP constituye el momento definido por la ley para el establecimiento de los hechos, previa valoración y decisión de la prueba, que se explicitará en una justificación con carga argumentativa creciente desde el veredicto a la sentencia definitiva.

Las limitaciones de tiempo y espacio de la deliberación, la falta de una estructura clara para llevarla a adelante, distorsiones en la forma de discutir y un exagerado acento en la motivación de los hechos como una cuestión que atañe sólo a la sentencia, pueden provocar deliberaciones deficientes o defectuosas. Esto puede traducirse en la generación de errores en el tratamiento de la prueba, en el establecimiento de los hechos y en una fundamentación incompleta del veredicto.

Por lo dicho, resulta útil revisar el modo en que se realiza esta labor y establecer ciertos presupuestos que permitan reforzar los beneficios esperables de una deliberación colegiada y epistemológicamente adecuada sobre los hechos.

Nuestra propuesta de trabajar guiados por una estructura básica con pasos metodológicos claros podría conducir a un fortalecimiento de dicho espacio institucional y permitir que la composición colegiada de miembros se materialice en decisiones sobre los hechos de mejor calidad que aquellas individualmente tomadas.

En tal sentido, proponer algunas guías específicas para mejorar la deliberación permitirá identificar dónde existen áreas de peligro de sesgos, mecanización o adopción de posiciones que se anteponen a los argumentos necesarios para sustentarlas, que transformen la deliberación en un trámite irrelevante o distorsionador y que sea la sentencia el espacio, ya no colegiado, que supla la falta o deficiencia del debate.

Lo anterior, tomando en cuenta que un tribunal formado por jueces profesionales no sólo está obligado a motivar sus decisiones y en específico sobre los hechos, sino que presentaría condiciones de entrenamiento en la argumentación probatoria que debiere diferenciarlo de otros colegiados que no tienen la misma carga de motivar, como el jurado.

Esta pretende ser una aproximación a una metodología de la deliberación sobre los hechos, para desarrollar, discutir y promover su reflexión. Si logramos generar una revisión de las propias prácticas por los jueces de tribunales colegiados, como el caso del TJOP chileno o similares de otras latitudes y promover el debate de esta temática, se cumplirá el propósito del trabajo.

BIBLIOGRAFÍA

- Accatino, D. (2003). La fundamentación de las sentencias. ¿Un rasgo distintivo de la judicatura moderna? *Revista de Derecho*, XV, 9-35.
- (2010). El modelo legal de justificación de los enunciados probatorios. En aa.vv, *Formación y valoración de la prueba en el proceso penal* (págs. 119-143). Legal Publishing Chile.
- (2011). Certezas, dudas y propuestas en torno al estándar de la prueba penal. *Revista de Derecho de la Pontificia Universidad Católica de Valparaíso*. XXXVII, segundo semestre, 483-511.
- (2019). Teoría de la prueba: ¿somos todos “racionalistas” ahora? *Revus*, 39, 85-102.
- Aguilera, E. (2022). Justificación epistémica, evidencialismo robusto y prueba jurídica. *Quaestio facti. Revista Internacional sobre Razonamiento Probatorio*, 3, 81-102.
- Aguiló, J. (2021). “El tribunal se retira a deliberar”. Un desafío teórico para juristas prácticos. *Revista Jurídica de les illes Balears*, 20, 11-30.

- Amaya, A. (2017). Virtudes y razonamiento probatorio. *Diálogos Jurídicos 2017. Anuario de la Facultad de Derecho de la Universidad de Oviedo*, 19-36.
- (2018). Ejemplaridad, virtud y autoridad judicial. *UNA Revista de Derecho (en línea)*, 3, 1-16.
- (2021). Virtudes, deliberación colectiva y razonamiento probatorio en el derecho. En J. Ferrer y C. Vázquez, *El razonamiento probatorio en el proceso judicial. Un encuentro entre diferentes tradiciones* (págs. 141-154). Marcial Pons.
- Anderson, T., Schum, D. y Twining, W. (2015). *Análisis de la prueba*. Marcial Pons.
- Araya, M. (2018). *Recurso de nulidad penal y control racional de la prueba*. Librotecnia.
- Aristegui, J. P. (2020). La prueba ilícita ante la bifurcación del tribunal penal. *Quaestio facti. Revista Internacional sobre Razonamiento Probatorio*, 1, 177-198.
- Atienza, M. (2005). *Las razones del Derecho. Teorías de la argumentación jurídica*. Universidad Nacional Autónoma de México.
- Bravo, S. (2022). *Prueba, valoración y decisión*. Editorial Librotecnia.
- (2023). Lenguaje, lógica y algunas repercusiones de la (in)definición del thema probandum. *Quaestio facti. Revista Internacional sobre Razonamiento Probatorio*, 4, 61-94.
- Bravo, S., Díaz, R. y Rettig, M. (Agosto-Diciembre de 2023). Siete pasos metodológicos para el establecimiento de los hechos en una decisión judicial. *Revista de Derecho Aplicado LLM UC*, 12. doi:https://doi.org/10.7764/rda.12.62617
- Calamandrei, P. (1960). *Proceso y democracia*. (Trad., H. Fix). Ediciones Jurídicas Europa-América.
- Carbonell, F. (julio de 2022). Un modelo de decisión judicial justificada para el proceso penal chileno. *Política criminal*, 17(33), 54-84.
- Cerda, R. (2003). *El Juicio Oral*. Metropolitana.
- (2016). *Juicio fáctico en las sentencias penales*. Librotecnia.
- Coloma, R. (2014). Dos es más que uno, pero menos que tres. El voto disidente en decisiones judiciales sometidas al estándar de prueba de la “duda razonable”. *Política criminal*, 9(18), 400-427.
- Damaska, M. (2015). *El derecho probatorio a la deriva*. Marcial Pons.
- Dei Vecchi, D. (Diciembre de 2020). Prueba libre, justificación epistémica y el noble sueño de los estándares de prueba. *Revista de Derecho (Valdivia)*, XXXIII, 25-48.
- Del Río, C. (2012). Estudio sobre el derecho al recurso en el proceso penal. *Estudios constitucionales*, 1, 245-288.
- Duce, M. y Riego, C. (2012). *Proceso Penal*. Editorial Jurídica de Chile.
- Ferrater Mora, J. (1964). *Diccionario de filosofía* (Quinta ed., Vol. I). Sudamericana.
- Ferrer, J. (2005). *Prueba y verdad en el derecho* (Segunda ed.). Marcial Pons.
- (2007). *La valoración racional de la prueba*. Marcial Pons.
- (2020). Sobre el deber de motivación de las decisiones probatorias y el juicio por jurados. La sentencia V.R.P., V.P.C. y otros vs. Nicaragua del CorteIDH. *Quaestio Facti. Revista Internacional sobre Razonamiento Probatorio*, 1, 359-382.
- (2021). *Prueba sin convicción*. Marcial Pons.
- Gajardo, T., & Hermosilla, F. (2020). *Curso de formación especializada para jueces y juezas de juzgado de garantía y de tribunal de juicio oral en lo penal*. Academia Judicial.
- Gama, R. (2011). El método de Wigmore y su aplicación en nuestra práctica jurídica. En X. Lluch, J. Picó i Junoy y M. González, *La prueba judicial. Desafíos en las jurisdicciones civil, penal, laboral y contencioso-administrativa* (pp. 163-179). La Ley.
- Gascón, M. y García, A. (2017). *La argumentación en el Derecho. Algunas cuestiones fundamentales*. Palestra.
- González Lagier, D. (2003a). Hechos y argumentos (Racionalidad epistémica y prueba de los hechos en el proceso penal) (I). *Jueces para la democracia*, 46, 17-26.
- (2003b). Hechos y argumentos (Racionalidad epistémica y prueba de los hechos en el proceso penal) (II). *Jueces para la democracia*, 47, 35-50.
- (2020). ¿Es posible formular un estándar de prueba preciso y objetivo? Algunas dudas desde un enfoque argumentativo de la prueba. *Revista Telemática de Filosofía del Derecho*, 23, 79-97.

- (2022a). Inferencia probatoria y valoración conjunta de la prueba. En J. Ferrer y S. C. Nación (Ed.), *Manual de Razonamiento Probatorio* (pp. 353-396). México.
- (2022b). Qué es el fundherentismo y qué puede aportar a la teoría de la prueba en el Derecho. En D. González Lagier, *Quaestio facti. Nuevos ensayos sobre prueba y filosofía* (Vol. II, pp. 87-111). Palestra.
- Haddad, P. (2017). Virtudes intelectuales y justificación epistémica en el ámbito probatorio. En A. y Amaya, *Emociones y virtudes en la argumentación jurídica* (págs. 37-58). Tirant Lo Blanch.
- Horvitz, M. I. y López, J. (2004). *Derecho Procesal Penal Chileno, tomo II*. Editorial Jurídica de Chile.
- Igartua, J. (1998). La dimensión dialéctica (y sus consecuencias) en la motivación de las sentencias. *Revista Vasca de Administración Pública*, 50, 175-2004.
- Kahneman, D., Sibony, O. y Sunstein, C. (2021). *Ruido. Un fallo en el juicio humano*. (Trad., J. Chamorro Mielke). Penguin Random House.
- Lo Guercio, N. (2018). Desacuerdos básicos entre pares epistémicos. *Ideas y Valores*, 168(67), 81-99.
- Martí, J. L. (2006). *La República Deliberativa. Una teoría de la democracia*. Marcial Pons.
- Maturana Baeza, J. (2014). *Sana crítica. Un sistema de valoración racional de la prueba*. Thomson Reuters.
- Maturana, C. y Montero, R. (2017). *Derecho Procesal Penal* (Tercera ed., Vol. II). Librotecnia.
- Posner, R. (2011). *Cómo deciden los jueces*. (Trad., V. Roca Pérez). Marcial Pons.
- Taruffo, M. (2008). *La Prueba*. Marcial Pons.
- (2010). *Simplemente la verdad - el juez y la construcción de los hechos*. Marcial Pons.
- Tuzet, G. (2021). *Filosofía de la prueba jurídica*. (Trad., D. Dei Vecchi). Marcial Pons.
- Vázquez, C. (Abril de 2018). La imparcialidad pericial y otras cuestiones afines. Confiabilidad, desacuerdos y sesgos de los expertos. *Isonomía*, 48, 69-107.
- Veleda, D. (2021). La decisión sobre la quaestio facti en los acuerdos de culpabilidad. *Quaestio Facti. Revista Internacional de Razonamiento Probatorio*, 2, 151-180.

LEGISLACIÓN Y JURISPRUDENCIA

Código penal (Chile) <https://bcn.cl/2fprj>

Código procesal penal (Chile) <https://bcn.cl/2f73h>

Código procesal penal (Costa Rica) https://www.pgrweb.go.cr/scij/Busqueda/Normativa/Normas/nrm_texto_completo.aspx?nValor1=1&nValor2=4129

Código procesal penal modelo para Iberoamérica

<https://biblioteca.cejamerica.org/bitstream/handle/2015/4215/textocodigoprocesalpenal.pdf?sequence>

Código procesal penal (Perú) <https://www.leyes.congreso.gob.pe/Documentos/DecretosLegislativos/00957.pdf>

España. Ley Orgánica 5/1995, de 22 de mayo, del Tribunal del Jurado (BOE núm. 122, 22/05/1995, p.15001-15021)

Sentencia de la Corte Suprema (Chile) 26152/2023 (Sala Penal, Segunda), de 28 de abril de 2023 <https://juris.pjud.cl/busqueda/u?coqji>

ESTÁNDAR DE PRUEBA EN CONTROVERSIAS POR OMISIÓN TRIBUTARIA: UNA MIRADA A LA LEGISLACIÓN Y JURISPRUDENCIA ECUATORIANA

Miguel Ángel Zumba Barbosa

Miembro del Foro de Abogados del Consejo de la Judicatura de Ecuador

Miembro del Instituto Ecuatoriano de Derecho Tributario

miguezumba1@gmail.com / mzumba@pbplaw.com

RESUMEN: En Ecuador, ningún litigio tributario se rige por un estándar de prueba. El planteamiento de este artículo es que, para que una hipótesis de omisión sea confirmada, deberán cumplirse las dos condiciones que exige el estándar 3 sugerido por Ferrer, es decir: (i) la hipótesis debe ser capaz de explicar los datos disponibles y predecir nuevos, con las pruebas que soporten esas predicciones; y, (ii) se debe refutar la hipótesis alternativa, si es plausible y compatible con la ausencia de responsabilidad del contribuyente, siempre que esté respaldada por pruebas que aporten algún grado de confirmación. De no superarse ambas condiciones, deberá prevalecer la ausencia de responsabilidad del sujeto tributario.

PALABRAS CLAVE: omisión; estándar de prueba elevado; contencioso-tributario.

STANDARD OF PROOF IN TAX OMISSION DISPUTES: AN OVERVIEW OF ECUADORIAN LEGISLATION AND CASE LAW

ABSTRACT: In Ecuador, no tax litigation is governed by a standard of proof. This article argues that, for a hypothesis of omission to be confirmed, the two conditions required by Ferrer standard 3 must be met, i.e.: (i) the hypothesis must be able to explain the available data and predict new

ones, with the evidence supporting those predictions; and, (ii) the alternative hypothesis must be rebutted, if it is plausible and compatible with the absence of taxpayer liability, as long as it is supported by evidence that provides some degree of confirmation. If both conditions are not met, the absence of liability of the taxpayer shall prevail.

KEYWORDS: omission; high standard of proof; tax litigation.

SUMARIO: 1. SUPUESTOS INICIALES.— 2. INTRODUCCIÓN.— 3. ¿CONCEPCIÓN SUBJETIVISTA O RACIONALISTA DE LA PRUEBA?.— 4. LA OMISIÓN TRIBUTARIA.— 5. ESTÁNDAR DE PRUEBA EN CONTROVERSIAS JUDICIALES DE OMISIÓN TRIBUTARIA: 5.1. Primer criterio: la facultad determinadora como poder de descubrimiento y corrección; 5.2. Segundo criterio: función de la jurisdicción contencioso-tributaria como instrumento en la garantía de derechos; 5.3. Tercer criterio: la carga de la prueba en materia tributaria y la presunción de legitimidad de los actos administrativos; 5.4. Cuarto criterio: problemas de diseño legislativo en la impugnación de las decisiones emanadas de la jurisdicción contencioso-tributaria; 5.5. Quinto criterio: costo social del proceso penal que resulta de una decisión adversa al contribuyente en la jurisdicción contencioso-tributaria.— 6. ¿CÓMO ADOPTAR EL ESTÁNDAR?.— 7. ¿CÓMO DEBERÍA APLICARSE EL ESTÁNDAR DE PRUEBA?: 7.1. Hipótesis de la autoridad tributaria; 7.2. Hipótesis del contribuyente; 7.3. Conclusión.— 8. BIBLIOGRAFÍA

1. SUPUESTOS INICIALES

Para el desarrollo de este documento, se asumirán determinadas tesis sobre las que no se profundizará pues no compete a esta investigación hacerlo. Estos supuestos iniciales son tres: (i) la aceptación de la concepción racionalista de la prueba (con todas sus consecuencias relevantes); (ii) su conveniencia por sobre la concepción subjetivista o persuasiva; y, (iii) la necesidad de contar con estándares de prueba para garantizar el derecho constitucional al debido proceso.

2. INTRODUCCIÓN

Las actividades de los particulares representan de forma creciente y acelerada un desafío para los entes fiscales, quienes se encuentran, cada vez más, frente a hechos y circunstancias económicamente relevantes que pueden dar lugar al nacimiento de tributos, lo cual deviene en que las disputas entre autoridades y sujetos tributarios sean de la más variada y diversa índole (Reyes y Páez, 2016, p. 110). En el marco de este amplio espectro de conflictividad, el trabajo se enfoca en los litigios donde se persigue la demostración de una omisión tributaria.

Por tal razón, este estudio sugiere la aplicación de un estándar probatorio elevado para este tipo de controversias, considerando la forma en que se fijan los hechos desde el procedimiento de determinación, la naturaleza de la jurisdicción contencioso-tributaria en la tutela de derechos, la regla de la carga de la prueba, la posibilidad

de impugnar la decisión que ponga fin al litigio y el coste social vinculado con la imputación penal. Todo esto en el marco normativo que brinda el ordenamiento jurídico ecuatoriano.

3. ¿CONCEPCIÓN SUBJETIVISTA O RACIONALISTA DE LA PRUEBA?

El Código Tributario (en adelante, CT) en su versión anterior a la norma hoy vigente, incorporaba la concepción subjetivista de la prueba en materia contencioso-tributaria, pues vinculaba el concepto de *verdad* con el *criterio judicial*. Según el ya derogado artículo 270, «a falta de prueba plena, el tribunal decidirá por las semipleonas, según el valor que tengan dentro del más amplio criterio judicial o de equidad». A renglón seguido, señalaba que las pruebas servirán para que los tribunales formulen presunciones, «siempre que con ellas pueda esclarecerse la verdad o ilustrar el criterio de los ministros».

El Código de Procedimiento Civil (en adelante, CPC), que estuvo vigente hasta el 2016, regulaba a la jurisdicción fiscal en todo lo que no estuviese expresamente normado en el Código Tributario, y también contenía referencias a la concepción subjetivista de la prueba. Por ejemplo, al referirse a la prueba testimonial, prescribía que el juez podía fundar su decisión en el testimonio de una persona aparentemente no idónea, «cuando tenga el convencimiento de que el testigo ha declarado la verdad». En la misma línea, establecía «en todo caso, en que el juez llegue al convencimiento de que se trata de una declaración falsa, suspenderá la diligencia, sin perjuicio de ordenar el enjuiciamiento penal correspondiente». Cuestión similar se advertía al referirse a la prueba pericial, pues el artículo 283 indicaba que «no es obligación del juez atenerse, contra su convicción, al juicio de los peritos». La noción de *convencimiento* y *convicción interna* estaba presente en la legislación y guiaba la noción de prueba en Ecuador.

Con la entrada en vigor del Código Orgánico General de Procesos (en adelante, COGEP) en el año 2016, la legislación ecuatoriana no cambió. La concepción subjetivista de la prueba volvió a incorporarse en la ley positiva. Según el art. 158, la finalidad de la prueba es «llevar a la o al juzgador al convencimiento de los hechos y circunstancias controvertidos». En el mismo sentido, el art. 172 prescribe que las mal llamadas presunciones judiciales adquieren relevancia cuando «conducen unívocamente a la o al juzgador al convencimiento de los hechos y circunstancias expuestos por las partes con respecto a los puntos controvertidos». Incluso el art. 164 establece que los juzgadores no tienen la obligación de emitir un pronunciamiento expreso relativo a la valoración de todos los medios probatorios practicados en juicio, sino solo respecto de aquellos «que le hayan servido para justificar su decisión».

Como no podía ser de otro modo, la doctrina procesal y la jurisprudencia tributaria han sido construidas en función de la íntima convicción del juzgador. Al referirse

al rol de la prueba en la jurisdicción contencioso-tributaria, Lascano Báez (2017, p. 200) indicó que su objetivo «no es otro que permitirle al tribunal resolver la controversia con arreglo a lo que éste, como cuerpo colegiado, considera que es verdad». Los Tribunales Distritales de lo Contencioso-Tributario (en adelante, TDCT) y la Sala Especializada de lo Contencioso-Tributario de la Corte Nacional de Justicia (en adelante, Corte de Casación y/o Sala Tributaria, indistintamente), también tienen una posición marcada en favor de la concepción subjetivista de la prueba.

A modo de referencia, el TDCT con sede en Quito, en el marco de un proceso en el que se discutía una impugnación por un ajuste de precios de transferencia, afirmó que la finalidad de la prueba pericial es la de «permitir el convencimiento del juzgador de los hechos respecto de los cuales debe aplicar la respectiva norma jurídica» (STDCT, juicio 17510-2023-00028, 23 de noviembre de 2023). El mismo tribunal, en otro proceso en el que se discutía la esencia económica de unas transacciones, señaló que la prueba «está ligada al interés que tiene cada parte de acreditar sus afirmaciones y lograr el convencimiento del juzgador [...]» (STDCT, juicio 17510-2022-00259, 13 de diciembre de 2023). De la misma manera, la Corte de Casación ha sostenido que los juzgadores de instancia sólo deben valorar las pruebas que «les llevaron a una convicción [...]». Esto implica que los jueces solamente deben «expresar en su resolución, la valoración de “todas las pruebas que le hayan servido para justificar su decisión” y no establece que el juzgador deba pronunciarse sobre todas y cada una de las pruebas que obran en el proceso» (SCNJ juicio 17510-2019-00389, dictada el 6 de octubre de 2022).

El breve diagnóstico de la legislación y jurisprudencia tributaria ecuatoriana da cuenta de dos elementos: (i) la íntima convicción judicial está arraigada en la tradición jurídica nacional, por lo que, hoy por hoy, resulta imposible controlar la racionalidad de las decisiones adoptadas sobre los hechos y las pruebas; y, (ii) aquello hace necesario pensar en el abandono definitivo de esta noción sobre la prueba, para acoger su concepción racionalista y así, construir estándares probatorios en controversias fiscales.

Sabiendo que la decisión sobre qué estándar fijar es política e idealmente debería ser adoptada por el legislador como principal órgano de representación y deliberación (Ferrer, 2021, pp. 149-150; Dworkin, 1985, p. 120), es necesario responder varias preguntas que parten sobre lo que, como sociedad, se espera y desea de la justicia tributaria: ¿Es necesario aplicar un estándar probatorio altamente exigente para evitar que aquellos contribuyentes que han satisfecho sus obligaciones tributarias oportunamente sean obligados a aportar en sobremedida al gasto público? O, por otro lado, ¿se debe pensar en un estándar menos exigente que permita condenar a más evasores? La respuesta no es pacífica.

Este trabajo se centra en examinar la aplicación de un estándar de prueba elevado en juicios tributarios donde se discuta la ocurrencia de una omisión en la declaración y pago de un tributo, para lo cual es preciso empezar por definir el concepto de omi-

sión tributaria según los aportes que ha hecho la jurisprudencia dictada por la Corte de Casación.

4. LA OMISIÓN TRIBUTARIA

¿Qué debe entenderse por omisión en el derecho tributario? Esta es la interrogante fundamental que ha suscitado amplio debate en la jurisprudencia ecuatoriana. ¿Cualquier disparidad entre lo declarado y lo determinado por la autoridad es, automáticamente, una omisión? ¿Las discrepancias sobre el contenido de un registro contable pueden catalogarse como omisión? ¿Qué supuestos se enmarcan en este gran paraguas conceptual? En esta sección se ofrecen ciertas directrices tendientes a aclarar este asunto.

El Diccionario panhispánico del español jurídico define a la omisión como la «conducta que consiste en la abstención de una actuación que constituye un deber legal». Narváez (2010, p. 210) describe a la omisión como la falta de intervención de un sujeto, destacando su relevancia en el ámbito legal en contraste con la comisión. Esta perspectiva conduce a la conclusión de que, desde un punto de vista legal, la omisión equivale a la abstención de realizar una acción, pero con implicaciones jurídicas significativas.

Al respecto, Barcena (2012, pp. 200 ss.) enseña que la omisión adquiere relevancia jurídica cuando un individuo deja de realizar una acción después de haber generado una expectativa, incluso sin infringir una obligación legal. Además, según el autor, la omisión adquiere trascendencia cuando, en el contexto de una actividad específica, no se cumplen medidas de seguridad propias de esa actividad, cuando el sujeto omite lo esperado o realiza algo diferente, o cuando lo hace a un nivel inferior al exigido por el derecho.

Tuzet (2013, p. 1) abona en esta discusión y señala que, para el campo jurídico, una abstención sólo puede identificarse en presencia de una norma que imponga una conducta específica, por lo tanto, el dejar de hacer algo, por sí solo, no tiene trascendencia jurídica, pero la adquiere cuando el sistema de normas vigente demanda una acción que ha sido desatendida.

El derecho tributario consagra una serie de deberes formales y sustanciales que requieren de los contribuyentes la ejecución positiva de un sinnúmero de acciones, de modo que, desde un punto de vista bastante genérico, dejar de hacerlas sería evidencia de una omisión. La doctrina parece entenderlo así, pues el ilícito tributario, concebido en un sentido amplio, es definido como «la vulneración de las normas tributarias mediante una conducta antijurídica, la cual puede consistir en la omisión de actos ordenados o en la ejecución de los prohibidos por la ley» (Ríos, 2005, p. 293).

Por un lado, los deberes formales reclaman el cumplimiento de determinadas prestaciones que, sin estar vinculadas con el pago de un tributo, son jurídicamente

trascendentes porque facilitan el control y supervisión de actividades económicas relevantes. Por ejemplo, es un deber formal solicitar a las autoridades tributarias las autorizaciones listadas en la ley, inscribirse en los registros pertinentes o presentar declaraciones de carácter informativo (CT, art. 96). De otro lado, los deberes sustanciales imponen sobre los sujetos pasivos la obligación de comunicar que ha tenido lugar el hecho imponible de un tributo, cuantificar la riqueza involucrada en ese acto y satisfacer la prestación pecuniaria que resulte de su determinación y liquidación (CT, arts. 16,18, 19 y 24). El objetivo que persigue la ley es que, si en la realidad ha tenido lugar el hipotético normativo que habilite el nacimiento de la obligación tributaria, este sea oportunamente informado y cuantificado.

Aunque la vulneración de deberes formales y materiales puede entenderse como un ilícito tributario, solo el incumplimiento de los segundos tiene la aptitud de activar la facultad determinadora de las autoridades tributarias, pues la transgresión de deberes formales tiene como consecuencia la aplicación de un gravamen punitivo, en ejercicio de la facultad sancionadora. Se trata de una competencia distinta, con connotaciones propias. Por ello, la *omisión tributaria* detectada en ejercicio de la facultad determinadora tiene lugar cuando se verifica el incumplimiento de obligaciones materiales, cuando se ha dejado de pagar un tributo debido. El supuesto al que se refiere este trabajo es justamente este y el análisis se enfoca en todos los aspectos que rodean su configuración.

La Sala Tributaria, a través de una serie de fallos, ha establecido los elementos constitutivos de la omisión en la declaración y pago de un tributo. En esencia, son dos: (i) la falta de reconocimiento de una operación que da nacimiento a una obligación tributaria; y, (ii) la voluntad de irrogar un perjuicio al fisco. Este segundo ha suscitado amplio debate.

Respecto al primero, la Sala Tributaria ha argumentado que la simple diferencia numérica entre el valor declarado y el monto glosado¹ por la autoridad fiscal no constituye una omisión. En otras palabras, si un contribuyente aduce que un casillero de su declaración es cero, mientras que la autoridad sostiene lo contrario, aquello no implica, al menos a primera vista, una omisión.

Al examinar procesos vinculados con el impuesto a la renta, la Corte de Casación ha dicho que hay omisión cuando el contribuyente, habiendo percibido ingresos gravados, deja de consignarlos, disminuyendo así la base imponible sobre la que se calcula su impuesto. Partiendo de esta consideración, es posible decir que se verifica esta con-

¹ «En el contexto del derecho tributario, la «glosa» se refiere al análisis detallado y crítico de una liquidación tributaria o declaración fiscal realizada por la autoridad fiscalizadora. Este proceso implica revisar minuciosamente los cálculos, las deducciones y los elementos utilizados para determinar la obligación tributaria de un contribuyente. La glosa tiene como objetivo identificar posibles errores, omisiones o discrepancias en la liquidación inicial, permitiendo a la autoridad fiscal corregir y ajustar la deuda tributaria correspondiente». (SCNJ, juicio 17510-2019-00311, de 2 de junio de 2023).

ducta abstencionista cuando el sujeto pasivo no consigna información que demuestra la ocurrencia del hecho imponible. Así, el criterio de la Sala Tributaria es que:

El lineamiento jurisprudencial antes referido advierte que, para que se considere que el plazo de caducidad de la facultad determinadora, sea de 6 y no de 3 años, debe demostrarse por parte de la autoridad de control, que el contribuyente *efectivamente ha ocultado un rubro que encaje en un tipo de renta o ingreso (partiendo del concepto de renta ingresos señalado en el Art. 2 de la LRTI) ya que no basta la simple diferencia numérica entre lo que pueda glosar la administración controladora y aquel valor declarado por el contribuyente*, pues de ser ese el caso, todas las determinaciones serían de 6 y no de 3 años y no se requeriría hacer diferencia alguna. (SCNJ, juicios 09501-2018-00126 de 20 de septiembre de 2021; 09501-2018-00175 de 20 de mayo de 2021; 09501-2018-00057 de 31 de mayo de 2021, 09501-2017-00791 de 21 de junio de 2021; 09501-2018-00281 de 30 de junio de 2021, 09501-2017-00736 de 22 de julio de 2021) (énfasis añadido).

Para ilustrar esta consigna, se expondrán un par de ejemplos extraídos de sentencias relativas al impuesto al valor agregado (en adelante, IVA) y el impuesto a la renta (en adelante, IR), en donde se rechazó la existencia de omisión debido a que, en realidad, existía una discrepancia de criterios entre los sujetos pasivos y la administración. Estas referencias buscan distinguir estos escenarios que no son semejantes pero que, de hecho, podrían llegar a confundirse con la omisión.

El Servicio de Rentas Internas, entidad encargada de administrar y recaudar los impuestos nacionales internos en Ecuador (en adelante, SRI), acusó a un contribuyente del sector de telecomunicaciones, de omitir la declaración y pago del IVA en las transferencias de equipos celulares en favor de sus clientes. Según decía la administración, la compañía solo declaró el IVA causado en la prestación del servicio de telefonía móvil, pues las facturas emitidas no segregaban el valor del bien y del servicio, lo que hacía imposible conocer si efectivamente se gravó con IVA las transferencias de bienes. El contribuyente adujo que su proceso de facturación estaba amparado en una norma reglamentaria que permitía, al prestar servicios y transferir bienes, facturarlos sin segregar el valor unitario por cada rubro, por lo que no existía omisión. La Sala Tributaria acogió la tesis de la compañía y dijo que no se verificó ninguna omisión porque las diferencias entre lo declarado y lo glosado tuvieron como origen el alcance, interpretación y aplicación de una norma jurídica (la norma reglamentaria invocada por el sujeto pasivo), de tal suerte que una discrepancia de criterios no puede asemejarse a una omisión².

En un proceso relativo al IR, las diferencias entre lo declarado por el contribuyente y lo determinado por la autoridad tributaria recaían en el contenido del casillero de gastos no deducibles. El contribuyente había registrado un gasto no deducible menor

² «6.3.6. La Sala, analizado el tema en controversia y en consonancia con el alcance del artículo 145 del RALRTI, norma reglamentaria que permite la facturación conjunta de equipo y servicio, se trata de la misma clase de servicio, es decir la misma actividad económica, por lo que, en concordancia con la línea jurisprudencial de la Sala, se trata de una diferencia de rubros que no justifica para considerar la existencia de ocultamiento de rubros alegada por la Administración Tributaria [...]» (SCNJ, juicio 17510-2020-00285, de 14 de noviembre de 2023).

del que exigía el SRI y, por ello, se afirmaba que la base imponible del impuesto se redujo, ocasionando un pago menor al realmente debido. La Sala señaló que, puesto que la diferencia se ocasionó en ese casillero, aunque el efecto del registro cuestionado sea la disminución de la base imponible, no podía hablarse de una omisión en sentido estricto, pues no se dejó de registrar ingresos gravados (el objeto mismo del impuesto)³.

A la luz de estos criterios, que solo son una muestra del gran desarrollo jurisprudencial sobre la materia, es plausible aducir que, uno de los requisitos para que se configure una omisión tributaria, requiere que, habiendo ocurrido el hecho generador del tributo, el contribuyente no lo reconozca en su declaración impositiva. Esta consigna automáticamente excluye del paraguas de la omisión a la discrepancia de criterios sobre normas tributarias, contables o de cualquier índole. Así, la prueba para acreditarla estará circunscrita a demostrar que ha tenido lugar el hecho de incidencia tributaria y que, habiéndose verificado, no se declaró por el sujeto pasivo.

Ahora bien, la Sala Tributaria también ha reconocido que el criterio anterior requiere complementarse con un insumo adicional: el ánimo deliberado de disminuir o anular la base imponible sobre la que se causa y paga el tributo. Este elemento *volitivo* sería indispensable en este tipo de controversias, pues es perfectamente plausible que un contribuyente cometa errores en su declaración sin que ello implique la pretensión de irrogar un perjuicio al fisco. La Corte de Casación razonó en el siguiente sentido:

Para esta Sala la apreciación de los elementos y circunstancias que rodean *la no inclusión de determinados ingresos en la declaración del contribuyente no necesariamente puede significar ocultamiento pues podría provenir de errores, omisiones, olvidos de buena fe que no evidencien un ánimo deliberado de disminuir los ingresos y por ende la base imponible y el tributo*. Elementos como la cuantía, el origen, la oportunidad de la percepción, pueden ser indicios de la existencia del ánimo deliberado de perjudicar al fisco y por ende no pueden ser considerados como ocultamiento. (SCNJ, juicios 09501-2017-00365 de 29 de junio de 2018; 17510-2017-00427 de 06 de mayo de 2019; 17510-2017-00528 de 11 de junio de 2016 y resolución 15-2009 de 18 de febrero de 2009) (énfasis añadido)

Vale la pena detenerse por un momento en este asunto. Si se acepta que el ánimo de perjudicar al fisco sea constitutivo de la omisión y, en consecuencia, un tema de prueba en un juicio contencioso-tributario, se estaría discutiendo un auténtico supuesto de defraudación tributaria, propio de la jurisdicción penal⁴. Por lo tanto,

³ «(iv) que la diferencia detectada en el acto administrativo impugnado no proviene de la ocultación de ingresos, sino por gastos no deducibles con afectación a la conciliación tributaria del impuesto a la renta correspondiente al año 2015, esto es, por incumplimiento de la normativa tributaria más no por ocultación de un rubro. 9.3.5. La Sala Especializada de lo Fiscal de la ex Corte Suprema de Justicia, mediante resolución ha señalado que «[...] una declaración se considera incompleta únicamente cuando existe ocultación de un rubro, y no cuando se evidencian omisiones o inexactitudes de rubros [...]» (SCNJ, juicio 175102019-00297, de 20 de mayo de 2022).

⁴ Doctrinariamente, los elementos constitutivos de la defraudación son: «a) engaño o aprovechamiento de errores; b) omisión de pagar un tributo de manera total o parcial; c) consecución de un bene-

resulta conveniente no incluir este elemento dentro del juicio contencioso-tributario, pues la esfera penal será la encargada de debatir en extenso al respecto. No sería eficiente debatir el mismo asunto en dos jurisdicciones distintas porque entonces no hay razón para distinguirlas. En un criterio reciente de la Corte de Casación, se confirmó esta posición, pues la decisión argumentó que no se requiere una comprobación del ánimo de irrogar un perjuicio al fisco para que se configure una omisión, por lo que, para validar la ocurrencia de esta conducta, basta con acreditar el ocultamiento de la operación que da nacimiento al hecho imponible:

6.3.7. Respecto a la falta de aplicación de los artículos 722 y 1475 del Código Civil, la Sala señala que al tenor del contenido literal de la norma, 94.2 del Código Tributario, los elementos que regulan tales disposiciones del Código Civil son ajenas a la aplicación de la caducidad, pues *la norma no condiciona a la intencionalidad expresa o manifiesta de perjudicar a la administración, pues será suficiente con demostrar la existencia del hecho del ocultamiento o la declaración parcial para la aplicación de este plazo de caducidad.* (SCNJ, juicio 17510-2020-00285, de 14 de noviembre de 2023) (énfasis añadido)

En definitiva, sobre la base de la jurisprudencia examinada, el concepto de omisión tributaria está atado a un supuesto en el que se verifica la disminución o anulación de la base imponible del tributo debido a la falta de reconocimiento de una operación que constituye un hecho impositivamente relevante, sea que haya mediado la intención de perjudicar al fisco o haya existido un error por parte del contribuyente. Para la jurisdicción contencioso-tributaria este factor resulta indistinto. A tal punto que, de confirmarse la omisión, proveniente del dolo o el error, la consecuencia jurídica sería la misma: el sujeto pasivo deberá satisfacer el gravamen tributario con los recargos e intereses prescritos en la ley.

A continuación, se indicará qué estándar de prueba se estima aplicable a esta clase de procesos y las razones que lo justifican.

5. ESTÁNDAR DE PRUEBA EN CONTROVERSIAS JUDICIALES DE OMISIÓN TRIBUTARIA

Tradicionalmente, la doctrina jurídica ha concebido la existencia de tres estándares de prueba. En el ámbito penal, se usa el estándar denominado «más allá de toda duda razonable», que resulta mucho más exigente que los estándares de «preponderancia de la prueba» o «prueba clara y convincente», que aparecen frecuentemente en la jurisprudencia emanada de la jurisdicción civil. Sin embargo, estos estándares presentan un problema fundamental: sus formulaciones son extremadamente vagas, con lo cual, dejan indeterminados sus criterios de aplicación.

ficio indebido; d) nexo de causalidad entre el engaño y el no pago del tributo u obtención del beneficio indebido». (Ríos, 2005, p. 332).

Ferrer construyó siete estándares probatorios que, en lugar de sustituir los tres referidos, concretan los lineamientos que pudieran ser perfectamente aplicables a cada uno de ellos. Esto quiere decir que sus propuestas de estándar podrían entenderse como «más allá de toda duda razonable», y buena parte del resto como «preponderancia de la prueba» o «prueba clara y convincente». Es necesario precisar, al igual que lo hace el autor, que estos estándares no son inmodificables, solo son una guía e ilustración de estándares metodológicamente bien formulados que pueden servir para la creación de otros.

Dicho esto, este artículo plantea que el estándar aplicable para la decisión final de un juicio contencioso-tributario donde se discute un supuesto de omisión es el que Ferrer explica bajo el número 3. Por ello, para que se tenga por probada una hipótesis sobre los hechos, será necesario una conjunción de las siguientes condiciones:

- a) La hipótesis debe ser capaz de explicar los datos disponibles, integrándolos de forma coherente, y las predicciones de nuevos datos que la hipótesis permita formular deben haber resultado confirmadas y aportadas como pruebas al proceso.
- b) Debe haberse refutado la hipótesis alternativa formulada por la defensa de la parte contraria, si es plausible, explicativa de los mismos datos y compatible con la inocencia del acusado/demandado o más beneficiosa para él, siempre que se haya aportado alguna prueba que le otorgue algún grado de confirmación. (Ferrer, 2021, p. 209).

Aplicando el estándar al supuesto de una omisión, este debería leerse del siguiente modo:

a) La hipótesis de la ocurrencia de una omisión tributaria debe ser capaz de explicar los datos disponibles, integrándose de forma coherente, y las predicciones de nuevos datos que la hipótesis permita formular deben haber resultado confirmadas y aportadas como pruebas al proceso.

b) Debe haberse refutado la hipótesis alternativa formulada por la defensa, es decir, la hipótesis de que la declaración presentada por el sujeto pasivo —de haberla— registró y cuantificó total y adecuadamente el hecho imponible, si es plausible, explicativa de los mismos datos y compatible con la ausencia de responsabilidad del contribuyente, o más beneficiosa para él, siempre que se haya aportado alguna prueba que le otorgue algún grado de confirmación.

La doctrina en el razonamiento probatorio ha sostenido que la decisión de elevar o reducir el umbral probatorio exigiría una evaluación de varios criterios no taxativos, entre ellos: (...) la gravedad del error en caso de condena falsa, el coste de error de la absolución falsa, las dificultades probatorias del tipo de casos, la incidencia de otras reglas que distribuyen el riesgo de error, la distribución de poder entre las partes y la etapa del proceso al que se aplique el estándar (Ferrer, 2021, p. 154 ss.). En el derecho administrativo, la decisión de elevar o bajar el umbral probatorio ha tomado en consideración, de forma bastante aleatoria y desorganizada, los principios que regulan una determinada materia —por ejemplo, aquellos que son propios del derecho ambiental o tributario—, el tipo de competencia y sus fines particulares —en su mayoría, la facultad sancionadora y el rol de la sanción administrativa en la sociedad—,

los bienes jurídicos afectados en el marco de esa competencia e incluso la protección en favor de una de las partes en conflicto (Letelier, 2018, pp. 219 ss.; Gaona, 2019, p. 28; Rebolledo, 2018, p. 13; Van Hasselt, 2015, pp. 229 ss.).

Los criterios que se exponen en este trabajo pretenden condensar la mayoría de las preocupaciones de la doctrina —individualizadas en el párrafo anterior— y trasladarlas al derecho tributario, concretamente, a las disputas donde esté en juego la ocurrencia de una omisión. Es por ello que, para justificar que el estándar debe ser elevado (número 3), en las siguientes secciones de esta investigación se explicará: (i) en qué consiste la facultad determinadora, cuál es su finalidad y cómo su dinámica de funcionamiento maximiza las probabilidades de formar conjuntos de evidencia cuantiosos y de calidad que sustenten la existencia de una obligación tributaria; (ii) la naturaleza de la jurisdicción contencioso-tributaria y su rol en la garantía de derechos como contrapeso del poder y de la confiscación; (iii) la regla de carga objetiva de la prueba que recae sobre el fisco, haciendo que, de no superarse el estándar, sea la parte procesal que soporte el peso de la insuficiencia probatoria; (iv) el deficiente diseño procesal que restringe la impugnación en materia de hechos; y, (v) los costos sociales asociados con la imputación penal —no los riesgos de condena—.

Estos factores son determinantes en la decisión del estándar aplicable, resultando en que aquel deba ser elevado.

5.1. Primer criterio: la facultad determinadora como poder de descubrimiento y corrección

La primera razón que justifica este estándar se extrae de las prerrogativas que confiere la facultad determinadora a las autoridades tributarias. En aplicación de esta potestad, cuya finalidad es garantizar el cumplimiento de las obligaciones tributarias materiales, el fisco tiene aptitudes casi exorbitantes para, en un tiempo razonable —de cuatro a seis años—, obtener de distintos agentes y por distintos medios, un cúmulo de información gigantesco que le permitirá acreditar la ocurrencia de una omisión. Un estándar de prueba menor sería admisible si las posibilidades de acceder a la información serían escasas o si el tiempo para ejercer la competencia fuese notablemente corto.

La obligación tributaria, en palabras de Saccone (2002, p. 81), «es el vínculo jurídico obligacional que se entabla entre el fisco como sujeto activo, que tiene la pretensión de una prestación pecuniaria a título de tributo, y un sujeto pasivo, que está obligado a la pretensión». Por supuesto, la obligación tributaria tiene particularidades. La más importante es que su fuente es la ley, por lo que el nexo jurídico entre sujeto activo y pasivo nace una vez que se verifica la ocurrencia de un hecho impositivamente relevante. Sobre esto, Valdez Costa (2016, p. 350) enseña que la ley debe recoger con precisión «la descripción del hecho gravado o elemento material de características variables que pueden agruparse en tres clases: posesión, obtención o utilización de una riqueza [...]».

La determinación es el mecanismo mediante el cual se hace saber a la autoridad fiscal que el hecho imponible ha tenido lugar y, en consecuencia, que luego de verificados los requisitos de ley, se entabla la obligación tributaria entre el particular y el Estado (Giuliani Fonrouge, 1997, p. 531).

En aras de detectar estos hechos de incidencia tributaria, los ordenamientos normativos prevén formas o sistemas que sirven para exteriorizar la ocurrencia del hecho imponible. Estos mecanismos se definen como sistemas de determinación, que recaen sobre el sujeto activo, pasivo o ambos. El primer caso tiene lugar cuando el tributo no requiere una declaración, pues el Estado, como ente acreedor del tributo, fija la cuantía de la obligación a pagar (CT, art. 90). El segundo surge con un ejercicio declarativo del contribuyente, quien informa la existencia del hecho generador y su cuantía (CT, art. 89). Finalmente, el tercero se conforma de una declaración del sujeto pasivo complementada con la intervención del sujeto activo (CT, art. 93).

Es particularmente importante la segunda forma de determinación, pues cuando el sujeto pasivo está obligado a reportar la existencia de una obligación tributaria, la declaración que presente le hará responsable de su contenido (CT, art. 101), siendo definitiva y vinculante para él. Consecuentemente, si esa declaración es incompleta, porque se ha omitido en todo o en parte un tributo debido, se gatillará responsabilidad a causa de esa omisión.

Conscientes de esta realidad, los ordenamientos jurídicos han dotado a las administraciones tributarias de una atribución que les permite verificar, por medio de un procedimiento reglado, que el ciudadano obligado a satisfacer un gravamen tributario lo haga en los términos y condiciones previstos en la ley, pudiendo, incluso, complementar o corregir las declaraciones impositivas objeto de revisión para que se ajusten a la realidad. Esta facultad se conoce como facultad de determinación (CT, arts. 96 y 94 de la sección 4 del Modelo de Código Tributario para América Latina), cuya finalidad es «asegurar la aplicación de los tributos de conformidad con lo previsto en la ley, esto es, que satisfagan sus tributos todos los llamados a ello por la ley y que lo hagan en la medida prevista por esta» (Piza, 2010, p. 625).

Por supuesto, el sistema sería muy ineficiente si, como resultado de la omisión que derive en una menor tributación, solo se condene al pago del tributo debido sin la imposición de un gravamen disuasorio (sanción) para prevenir ilícitos futuros. En Ecuador, este gravamen viene en la figura de un recargo que alcanza el 20 por 100 del impuesto determinado y que, básicamente se explica como un «castigo» resultante del ejercicio de la competencia (CT, art. 90 y Resolución dictada por la Corte Nacional de Justicia el 4 de mayo de 2011).

Ahora bien, en el marco de esta competencia que, según Gonzáles (2000, p. 22), comprende los poderes de «investigación y comprobación» de la obligación tributaria, la autoridad fiscal tiene prerrogativas extensas para determinar la existencia de un hecho imponible (Troya, 2014, p. 255 y Folco, 2011, p. 485). La actividad probatoria en el contexto de este procedimiento tributario adquiere gran importancia,

pues permite que la autoridad verifique y constate los «hechos susceptibles de dar lugar al nacimiento de una obligación tributaria o de las circunstancias que inciden en la determinación de su cuantía» todo lo cual «constituye un paso necesario para la aplicación de las normas que determinan la concreta obligación tributaria [...]» (Piza, 2010, p. 553).

Cuando la determinación tiene como base la revisión de una declaración, el sujeto pasivo debe mostrar su contabilidad, registros y todos los documentos que soporten esos registros. Además, la administración se podrá valer de sus propias bases de datos, las bases de datos de instituciones públicas o privadas y cualquier documento en poder de terceros que esté vinculado con la determinación (CT, arts. 91 y 91.1). La finalidad de la ley es que el fisco ejerza una actividad de indagación exhaustiva, para lo cual, puede utilizar documentación formal y material: llámese contratos, informes y todo cuanto pudiese demostrar la ocurrencia del hecho que motiva la fiscalización⁵.

Si no existe declaración, el fisco conserva opciones para escrutar y controlar la ocurrencia del hecho impositivamente relevante. Aun en este escenario, que le es naturalmente desfavorable, la autoridad tributaria puede utilizar y requerir información de terceros, datos de cualquier sujeto pasivo, incluso, recurrir a la verificación de hechos, circunstancias y cualquier elemento que le permita establecer la configuración del hecho generador y su cuantía (CT, arts. 92 y 93). En cualquier escenario, la autoridad tributaria conserva prerrogativas de investigación extensas que no dependen del sujeto investigado.

A esto hay que sumar el cúmulo de información accesible en virtud de acuerdos internacionales. Ecuador tiene suscrito 23 tratados internacionales para evitar la doble imposición con países como Suiza, Uruguay, España, Francia, etcétera⁶. Además, desde el 26 de abril de 2017, Ecuador es miembro del Foro Global, que cuenta con la participación de 150 miembros⁷. Con ello, la autoridad tributaria tiene una puerta que le permite acceder a datos de administrados a nivel internacional, ampliando estrepitosamente la posibilidad de controlar y determinar los hechos de incidencia tributaria.

Tan abrumadora es la facultad determinadora que parece dar prerrogativas ilimitadas. Justamente por esa razón la doctrina ha advertido que se requieren límites materiales claros que, de sobrepasarse, anulan el ejercicio de la competencia. Estos límites pueden recaer sobre el objeto mismo de la auditoría y sobre el tiempo para ejercerla.

⁵ Al respecto, puede tomarse, con pinzas, el precedente jurisprudencial obligatorio 07-2016 emitido por el Pleno de la CNJ, el cual señala que, para soportar la esencia de una transacción a efectos de considerar la existencia de un gasto deducible del impuesto a la renta, se debe verificar documentos «fuente» de la obligación, sin establecer de forma taxativa cuáles serían esos documentos, lo cual, desde mi perspectiva es acertado.

⁶ Información extraída del siguiente sitio web: <https://www.sri.gob.ec/fiscalidad-internacional2>

⁷ Información extraída del siguiente sitio web: <https://www.sri.gob.ec/foro-global>

En cuanto a lo primero, la administración solo podría indagar sobre los datos de incidencia tributaria y respecto de las situaciones que motivan la auditoría (Rodríguez Bereijo, 2007, p. 119 y Pino, 2023, pp. 100 ss.). En lo que concierne a lo segundo, es decir, la temporalidad de la competencia, resulta evidente que, en un estado constitucional, un poder como este no puede concebirse sin la existencia de un plazo fatal. La caducidad viene a ser el medio para este fin debido a que es una institución que precautela la seguridad jurídica pues fija puntos temporales claros que, en caso de ser superados, impiden que las autoridades tributarias ejerzan válidamente la potestad fiscalizadora sobre los particulares (Rodríguez Russo, 2012, p. 63). Acorde a lo prescrito en el artículo 94 del CT, toda autoridad tributaria puede ejercer válidamente su facultad determinadora dentro de cuatro años si el contribuyente presenta su declaración; o seis años si no lo hace.

Con un poder como este, que fácilmente puede llevar hacia la arbitrariedad y las desviaciones de poder, la autoridad tributaria está en la aptitud de nutrirse con un cúmulo gigantesco de datos que le permitirá evaluar y cuantificar un hecho de incidencia tributaria que ha sido omitido. El espacio temporal y la cantidad de información a la que el Estado tiene acceso es abrumadora y este es un factor determinante para la propuesta del estándar. Es mucho mejor tomar decisiones sobre la base de un gran volumen de información que puede recopilarse en un tiempo razonable porque reduce las posibilidades de que un evasor sea absuelto y un cumplidor sea condenado⁸. Ciertamente, si la situación fuera diferente, esto es, que el plazo para ejercer la competencia sea corto y que exista imposibilidad de acceder a datos relevantes, podría pensarse en reducir el estándar pues sería irrazonable exigir la satisfacción de un umbral elevado sin que existan herramientas que permitan alcanzarlo. Como ocurre justamente lo contrario, y se entiende que una competencia tan sensible se ejerce con prolijidad, hay una expectativa razonable de que la evidencia de respaldo de la hipótesis de omisión sea abundante y de calidad.

5.2. Segundo criterio: función de la jurisdicción contencioso-tributaria como instrumento en la garantía de derechos

La segunda razón que justifica la aplicación del estándar probatorio sugerido se extrae de las características intrínsecas de la jurisdicción contencioso-tributaria. En su concepto, los magistrados que pertenecen a esta jurisdicción son un contrapeso del poder y deben procurar el resguardo de los derechos de los sujetos pasivos, evitando con celo la confiscación por medio de conductas impositivas, aparentemente legítimas. Es por ello que una de las consecuencias relevantes de esta protección debe traducirse en un estándar probatorio elevado que reduzca las probabilidades de un mayor número de falsos condenados a causa de una omisión tributaria.

⁸ *Vid.* Ferrer (2021, p. 122).

Según Troya (2014, pp. 319-321), un juicio tributario es un híbrido entre el proceso civil y el proceso penal. Lo que le da esta naturaleza es el control que ejercen los jueces tributarios sobre los actos impugnados y sus antecedentes, a razón de la importancia pública que revisten los asuntos fiscales. En la misma línea, autores como Neira (2016, pp. 241-244) y Simone (2017, pp. 459-460) señalan que los juzgadores no solo deben resolver el conflicto planteado, además, deben revisar «la actuación administrativa, lo que puede suponer ir más allá del controvertido, claro está, sin desatender lo peticionado por las partes y sin atribuirse competencias que no le han sido asignadas».

En Ecuador, el legislador reconoció normativamente este carácter híbrido del proceso contencioso-tributario, pues estableció que los jueces en esta materia deben ejercer un control de legalidad de los actos sujetos a su conocimiento, con un componente adicional: la tutela de derechos en el marco de la relación jurídico-tributaria. Según lo dispone el COGEP, esta jurisdicción tiene por objeto:

tutelar los derechos de toda persona y realizar el control de legalidad de los hechos, actos administrativos o contratos del sector público sujetos al derecho tributario o al derecho administrativo; así como, conocer y resolver los diversos aspectos de la relación jurídico-tributaria o jurídico administrativa, incluso la desviación de poder. (art. 300).

Por ello, al resolver una controversia sometida a su conocimiento, los jueces decidirán

[...] con claridad los puntos sobre los que se produjo la controversia y aquellos que en relación directa a los mismos comporten control de legalidad de los antecedentes o fundamentos de la resolución o acto impugnados, supliendo incluso las omisiones en que incurran las partes sobre puntos de derecho, o se aparte del criterio que aquellas atribuyan a los hechos. (COGEP, art. 313).

Claramente, la jurisdicción contencioso-tributaria concebida en los términos previstos por el legislador ecuatoriano, tiene un fin estrictamente proteccionista, pues busca tutelar que toda actuación administrativa esté apegada a derecho y respete las garantías básicas del contribuyente. Este rol garantista de derechos ha sido reconocido por la Corte Nacional de Justicia en su jurisprudencia más reciente.

La Sala Especializada de lo Contencioso-Administrativo de la Corte Nacional de Justicia (en adelante, Sala Administrativa) ha reiterado que el control de legalidad exige corroborar el entendimiento correcto del ordenamiento jurídico, supliendo los defectos en derecho de las partes procesales en aras de garantizar una justicia efectiva (SCNJ, juicios 01803-2019-00258 de 21 de septiembre de 2020; 17811-2013-0648 de 10 de febrero de 2022; 13802-2017-00131 de 27 de noviembre de 2018 y 09802-2020-00424 de 19 de septiembre de 2022). Por su parte, la Sala Tributaria de la CNJ tiene el criterio de que el papel de la justicia tributaria es ser garante de los derechos de los ciudadanos, más allá del tradicional control de legalidad (SCNJ, juicio 17510-2018-00418, de 25 de octubre de 2022 y juicio 17510-2018-00383, de 17 de febrero de 2023).

Una de las razones, quizá la principal, por las que esta jurisdicción está investida con el control de juridicidad para garantizar los derechos de los particulares, es porque la imposición tributaria no puede desprender al ciudadano de su patrimonio y anular su núcleo esencial, convirtiéndose así en una práctica confiscatoria⁹, tal como ha reiterado la Corte Constitucional del Ecuador (en adelante, Corte Constitucional) al establecer el alcance del principio constitucional de no confiscatoriedad y capacidad contributiva.

La protección en contra de la confiscación como resultado de la imposición tributaria está justificada y reforzada, a diferencia de lo que ocurre en otras materias donde también pugna el derecho de propiedad, al menos, por dos razones. La primera irradia a partir del resultado del proceso de fiscalización. Como se indicó, los procesos de fiscalización que resulten en la confirmación de la hipótesis de omisión impondrán sobre el sujeto tributario la obligación de pagar un tributo junto con un gravamen de naturaleza sancionatoria. Esto hace que, al condenar erróneamente al cumplidor tributario, se le exija un desprendimiento patrimonial injustificado que deviene de la satisfacción de un tributo no debido, junto con una sanción pecuniaria adicional que alcanzará el 20 por 100 de la obligación principal. En definitiva, el resultado del error sería la legitimación de la confiscación —por medio de la facultad determinadora— y el ejercicio arbitrario del *ius puniendi* estatal —por medio del recargo—.

Y la segunda justificación parte de que la disputa tributaria trae consigo una balanza desequilibrada favorablemente al Estado. En función de lo dicho en la sección anterior de este artículo, debe recordarse que el fisco tiene recursos y prerrogativas normativas exorbitantes para detectar los hechos de incidencia tributaria que lo sitúan en una posición siempre cercana con la desviación del poder, de allí que exista una construcción doctrinaria muy fuerte respecto a las limitaciones materiales de sus competencias. Si esta jurisdicción no tuviese la posibilidad de controlar la legalidad de las actuaciones con miras a garantizar los derechos de los ciudadanos, entonces no existirían mecanismos que equilibren las desigualdades entre el fisco y los contribuyentes, volviendo inoficioso al sistema procesal¹⁰.

⁹ Por ejemplo, los fallos: SCC, 46-18-IN/23, de 6 de septiembre de 2023; SCC, 65-17-IN/21, de 19 de mayo de 2021; SCC, 47-15-IN/21, 10 de marzo de 2021; SCC, 10-21-IN/23, de 22 de noviembre de 2023. Otros tribunales constitucionales también han emitido criterios respecto a la prohibición de confiscación en materia tributaria, por ejemplo, el Tribunal Constitucional de Perú, en las Sentencias Nos. 0041-2004-AI/TC, de 11 de noviembre de 2004, y 00053-2004-PI/TC, de 16 de mayo de 2005. Igualmente, la Corte Suprema de Argentina en el Fallo 186-170 y la Corte Constitucional de Colombia en la Sentencia C-293/20.

¹⁰ Al respecto Rozen- Zvi y Fisher (2008, pp. 91-92), advierten que para que el sistema procesal adversarial sea funcional, se requiere una situación de igualdad entre las partes, pues de lo contrario, la balanza se inclinará hacia la que tenga más poder y recursos. Al referirse a las distinciones entre un litigio contra el Estado en contraste con uno en que la contraparte es un particular, indican que existen varias herramientas que inclinan la balanza en favor del Estado, entre ellas se encuentra el mismo diseño procesal, fijado por el mismo Estado, los recursos, funcionarios y el monopolio del uso de la fuerza, entre otros.

De este modo, la jurisdicción contencioso-tributaria y el principal bien jurídico tutelado, este es, la protección jurídica al derecho de propiedad frente a la confiscación estatal debe traducirse en un estándar de prueba exigente como una herramienta adicional para equiparar las posiciones de los actores principales en juego. La necesidad de superar el umbral contenido en el estándar 3 garantiza que la jurisdicción tributaria alcance los fines perseguidos por el sistema jurídico pues se constituye como un filtro que impedirá que el Estado, quien por regla general está en una posición de ventaja frente al particular, castigue a un ciudadano a través de prácticas aparentemente legítimas —ejercicio de la facultad determinadora— con la confiscación de todo o parte de su patrimonio para satisfacer una carga impositiva que no le corresponde, y le imponga una sanción que también es de naturaleza patrimonial.

Todo lo dicho no necesariamente implica que, en controversias ajenas a la omisión, esta protección reforzada en contra de la confiscatoriedad pueda relativizarse y, con ello, reducir el estándar de prueba con miras a maximizar la vigencia de otros principios, por ejemplo, el de generalidad. Esto podría ocurrir en varias controversias, especialmente aquellas donde el objeto en discusión sea la aplicación de cláusulas anti-elusión para gravar a contribuyentes que han abusado o han cometido fraude a la ley con la sola finalidad de conseguir beneficios tributarios o reducir su carga fiscal. En esos casos, donde las dificultades probatorias son elevadas y está en juego la confianza social de que todos pagan impuestos en su justa medida, sería admisible evaluar una reducción del estándar para precautelar dicha finalidad, sin que ello implique desnaturalizar esta jurisdicción. Esta misma reducción del estándar estaría justificada en controversias donde el único resultado sea la imposición de sanciones pecuniarias de bajas cuantías, o donde el cierre del establecimiento o la suspensión de actividades sea menor y no incida significativamente en el patrimonio de los contribuyentes.

En fin, hay un amplio catálogo de situaciones por explorar, pero la que compete a esta investigación, aquella vinculada con la omisión, requiere de un estándar probatorio exigente por las razones expuestas.

5.3. Tercer criterio: la carga de la prueba en materia tributaria y la presunción de legitimidad de los actos administrativos

Otra de las razones que justifica la elección de un estándar probatorio elevado es la carga de la prueba objetiva en materia tributaria. Muchos tribunales entienden, incorrectamente, que el fisco estaría libre de cargas probatorias debido a la presunción de legitimidad de los actos administrativos. Por esa razón, a falta de prueba, deciden en detrimento de los contribuyentes. En esta sección se examinará de donde surge esta confusión, clarificando que son las administraciones tributarias las que deben soportar las consecuencias adversas por no haber acreditado los enunciados fácticos que justifican la imputación de la omisión tributaria.

El art. 169 del COGEP establece que «es obligación de la parte actora probar los hechos que ha propuesto afirmativamente en la demanda y que ha negado la parte demandada en su contestación». Sin embargo, existe una regla especial que rige para los juicios contencioso-tributarios. El segundo inciso del art. 311 de la misma ley establece que «con respecto a los actos tributarios impugnados, corresponderá a la administración la prueba de los hechos o actos del contribuyente, de los que concluya la existencia de la obligación tributaria y su cuantía». La jurisprudencia histórica y reciente de la Sala Tributaria ha ratificado la vigencia de esta norma respecto a las autoridades fiscales, aunque no con consistencia.

La Corte de Casación dijo que «si la administración tributaria estableció la existencia de hechos imponible distintos a la actividad de la construcción (rendimientos financieros) aplicando el art. 259¹¹ del Código Tributario, correspondía su demostración a efectos de ejercer su facultad determinadora [...]» (SCNJ; juicio 17510-2016-00192, de 27 de abril de 2023). Del mismo modo, ha afirmado que «cuando la Administración Tributaria en el ejercicio de la facultad determinadora, incluya ingresos no declarados por el contribuyente, le corresponde la carga de la prueba, al tenor de lo previsto en el artículo 311, inciso segundo del COGEP» (SCNJ; juicio 17510-2018-00167, de 29 de septiembre de 2022)¹².

Pese a que la legislación procesal establece que las autoridades tributarias tienen la carga de probar los hechos o actos de los que concluyan la existencia de una obligación tributaria, esto no se cumple, o al menos, no con uniformidad. El factor que incide directamente en ello es la existencia de la mal llamada presunción de legitimidad que cobija a toda actuación administrativa.

Según esta «presunción», todo acto de la administración pública se entiende legítimo —en el sentido de que ha sido emitido conforme a derecho— hasta que una autoridad judicial declare lo contrario (Guerrero, 2019, p. 273 y Vidal, 1985, p. 258). El art. 82 del CT así lo indica al establecer que «los actos administrativos tributarios gozarán de las presunciones de legitimidad y ejecutoriedad y están llamados a cumplirse; pero serán ejecutivos, desde que se encuentren firmes o se hallen ejecutoriados».

¹¹ Actual 311 del COGEP.

¹² En el mismo sentido: SCNJ, dictada en el juicio 17510201900004, de 20 de agosto de 2021: «La Sala deja sentado que, en aplicación de lo preceptuado en el art. 311 del COGEP, a la Administración Tributaria le corresponde probar los cargos que pretende establecer en contra del contribuyente, cuya declaración no ha sido objetada. Por estas consideraciones, se desvanece la glosa». Pronunciamientos anteriores en este sentido, por ejemplo, SCNJ, emitida en el recurso de casación No. 129/93, de 23 de marzo de 1994: «[...] si bien es verdad que es obligación del actor «probar los hechos que ha propuesto afirmativamente en la demanda y que ha negado expresamente la Autoridad demandada, salvo aquellos que se presuman legalmente...», el Código Tributario hace la excepción en el último inciso del Art. 274, respecto de los actos impugnados, que corresponderá a la Administración la prueba de los hechos y actos del contribuyente, de los que se concluya la existencia de la obligación tributaria y su cuantía: lo cual constituye un deber sustancial de la Administración Tributaria [...]». Asimismo, STDCT, dictada en el juicio 17504-2011-011, de 25 de junio de 2020.

Sin embargo, en lo absoluto esta regla jurídica recoge una presunción propiamente dicha y, por tanto, de ningún modo «invierte» las cargas probatorias en detrimento del contribuyente, tal como sostiene buena parte de la doctrina del derecho administrativo¹³ y algunos fallos de la Sala Tributaria¹⁴ —de forma contradictoria con aquellos donde reconoce que la carga de la prueba objetiva recae sobre el fisco—. Si se parte de que la presunción es una especie de argumento según el cual ciertos hechos base se vinculan con determinados enunciados fácticos a través de una garantía que justifique la aceptación del enunciado (Dei Vecchi, 2019, p. 6), entonces no es posible aceptar que la «presunción de legitimidad» realmente lo sea. No existe un hecho base y un enunciado fáctico que resulte del primero. El único hecho que existe es la emisión del acto administrativo.

En tal sentido, el efecto de esta regla es permitir que un acto administrativo continúe desprendiendo efectos siempre que no sea eliminado del sistema jurídico por una decisión judicial. Pero ello no es una consecuencia propia del acto administrativo. Lo mismo ocurre con la mayoría de las normas que conforman el ordenamiento jurídico: si aquellas no son declaradas ilegales o inconstitucionales, simplemente se mantendrán vigentes y deberán cumplirse. Por ende, es incorrecto afirmar que cuando se impugna una actuación del poder público se invierte la carga de prueba sobre el ciudadano. Si esto fuera así, entonces el art. 311 del COGEP estaría virtualmente derogado y no existiría ninguna razón para que la norma subsista en el ordenamiento jurídico.

Sobre esto, Mejía (2013, pp. 66-68) afirma que cuando se impugna un acto administrativo de carácter tributario, la autoridad queda en la posición de probar los reparos que le hayan motivado a decidir que el proceder del contribuyente fue errado. Para el autor, solo así se garantiza la *presunción de corrección*¹⁵ que asiste a todo sujeto pasivo, el cual es una derivación de la presunción constitucional de inocencia:

¹³ Así lo sostiene, por ejemplo, Pérez Camacho (2021, p. 90) quien señala «[l]a referencia a la «presunción» de legitimidad le atribuye el efecto corriente de las presunciones, que es de invertir la carga de la prueba, es decir que los actos se presumen legítimos mientras no se pruebe lo contrario».

¹⁴ La misma Sala Tributaria ha reconocido también en su jurisprudencia la vigencia de la presunción de legitimidad de los actos tributarios, atribuyendo, incorrectamente, el efecto de inversión en la carga de la prueba: «Claramente se identifica en la normativa que el actor está obligado a probar los hechos, para hacer valer la pretensión en su demanda [...] Al no haber desvirtuado el contribuyente que los actos impugnados son ilegales y que no cumplieron con lo que dispone la normativa al respecto, esta Sala no puede pronunciarse sobre el objeto de la litis y considera que la Resolución Impugnada N° 109012010RREC017391 está revestida de la presunción de legitimidad y ejecutoriedad de conformidad a lo que establece el Art. 82 del Código Tributario». (SCNJ, recurso No. 383-2011, de 31 de octubre de 2013).

¹⁵ STDCT dictada en el juicio 17510-2023-00028 de 27 de noviembre de 2023. En lo pertinente: «[...] busca confrontar objetivamente la información contenida en la determinación efectuada por el contribuyente, a través de su declaración —la cual se presume veraz— para comprobar la correcta aplicación de las normas constitucionales, legales, reglamentarias y contables a las que se somete el régimen jurídico tributario, debiendo generar el procedimiento, certidumbre y seguridad sobre los resultados que arroje».

La dilución de la presunción de validez del acto administrativo tributario que hubiese sido impugnado tiene como consecuencia que las afirmaciones de la Administración Tributaria contenidas en el acto queden en tela de juicio al igual que las del actor sujeto pasivo y, por tanto, deban ser forzosamente probadas ante el juez. Esto que deviene de toda lógica, es recogido expresamente por el segundo inciso del citado artículo 259 del Código Tributario [actual 311 del COGEP] [...].

En este escenario, es claro que son las administraciones las que deben acreditar la existencia del hecho generador que ha sido omitido, por lo que las consecuencias de no hacerlo las perjudicarán. Esta conclusión se sustenta no solamente en el marco normativo que ha sido explicado, sino también en cómo opera el proceso de determinación. Conforme se refirió en este artículo, es la propia entidad la que, en ejercicio de su facultad determinadora, está obligada a construir un expediente administrativo que cuente con todos los elementos que le hayan llevado a sostener una hipótesis fáctica: en este caso, la hipótesis de una omisión. No hacerlo así afectaría la voluntad administrativa porque el resultado de la auditoría no estaría respaldado en prueba.

La dinámica del sistema procesal ratifica este argumento ya que la ley le exige a la entidad fiscal acompañar el expediente íntegro de la determinación junto con la contestación a la demanda que se sigue en su contra. Aunque la legislación procesal vigente no lo dice, la consecuencia jurídica de no entregar el expediente con todas las pruebas que sirvieron para formular la hipótesis fáctica es la pérdida del proceso por falta de prueba, porque la autoridad no tendrá evidencia que justifique su planteamiento en juicio y, además, porque el órgano jurisdiccional se vería impedido de efectuar el control de juridicidad previsto en la ley. En el Recurso 339-2010, de 1 de octubre de 2010, y la Resolución 119-2010 de 19 de abril de 2010, sustanciados por la derogada Ley de la Jurisdicción Contencioso-Administrativa, la Sala Administrativa así lo confirmó.

De este modo, si bien el expediente no es la única evidencia en juicio, sin duda es la más importante, porque el plazo que se tiene para ejercer la facultad de determinación es amplio, mientras que, el plazo para ejercer el derecho de contradicción al contestar la demanda es de tan solo treinta días, en los cuales la posibilidad de abonar con información adicional a la contenida en el expediente es ínfima y limitada, pues tampoco es admisible una motivación extemporánea del acto administrativo.

Por lo tanto, la intención y la literalidad de la ley es que sea la autoridad la que acredite la hipótesis de omisión, no el contribuyente. Sustentar lo contrario sobre la única base de la presunción de legitimidad, no es compatible con los pilares que fundan a la facultad determinadora y la carga que ha impuesto el sistema jurídico sobre el ente fiscal.

Así, si un sujeto pasivo impugna una actuación administrativa confirmatoria de una hipótesis de omisión, será la autoridad quien deberá acreditar su ocurrencia, con el riesgo de que, de no hacerlo, pierda el proceso. Por otro lado, si el contribuyente ha ejercido su derecho a la defensa desde la fase administrativa hasta la fase judicial realizando afirmaciones para que se aplique una determinada norma en su favor —como

habitualmente ocurre—, deberá acreditar esas afirmaciones con prueba, pero el estándar que se debe considerar para este supuesto es correlativamente inferior, nunca podría ser equiparable al de culpabilidad¹⁶. Por lo tanto, desde el estándar 3 sugerido para las controversias de omisión, se requiere, necesariamente, que el sujeto pasivo haya aportado material probatorio para el análisis, caso contrario, el órgano jurisdiccional podría prescindir del descarte de la hipótesis alternativa.

5.4. Cuarto criterio: problemas de diseño legislativo en la impugnación de las decisiones emanadas de la jurisdicción contencioso-tributaria

La cuarta razón que justifica la aplicación del estándar probatorio sugerido es consecuencia de la falta de una segunda instancia para controlar de modo amplio la racionalidad de las inferencias probatorias. En Ecuador, el único medio de impugnación vertical para un juicio tributario es el recurso de casación. Los problemas del diseño legislativo de este recurso hacen que una decisión errada sobre los hechos sea prácticamente inmodificable, y esa consideración exige un estándar probatorio elevado como un mecanismo de prevención ante las condenas falsas, es decir, la condena por omisión en perjuicio de un sujeto pasivo que ha cumplido con sus obligaciones tributarias.

En Ecuador no está previsto el recurso ordinario de apelación para la jurisdicción contencioso-tributaria, como si lo está para otras materias. Esto hace que el único medio de impugnación previsto para refutar una decisión en esta jurisdicción sea el recurso extraordinario de casación. Este diseño legislativo, fuertemente criticado porque anularía la garantía constitucional del doble conforme (Aguirre *et al.*, 2022, pp. 13-17) complica, pero no anula la posibilidad de recurrir una sentencia de instancia, en lo que concierne al control de las inferencias probatorias. Sin embargo, la falta de un estándar de prueba es un factor determinante que incide y traba aún más al control casacional.

Conforme a la legislación ecuatoriana, para que los hechos que han sido declarados como probados por un tribunal sean susceptibles de cuestionamiento a través del recurso de casación, deben concurrir tres condiciones: (i) que la declaración de esos hechos sea una consecuencia de una infracción a un precepto jurídico de valoración probatoria¹⁷; (ii) que ese vicio derive, como un efecto «carambola», en la vulneración de una

¹⁶ Al respecto, esta idea también se desprende del criterio de Ferrer Beltrán (2021, p. 156), quien sostiene que el estándar de prueba de culpabilidad y de inocencia en procesos penales debería ser distinto. Esto, aplicado al caso concreto me permite sostener que el estándar de prueba para declarar la responsabilidad del contribuyente frente a su inocencia debe ser diferente.

¹⁷ La jurisprudencia nacional en materia tributaria ha sido muy consistente en señalar que la norma de valoración probatoria prevista en nuestra legislación es el art. 164 del COGEP, cuyo segundo inciso reconoce que la prueba debe valorarse siguiendo las reglas de la sana crítica y que, además, debe ser apreciada en su conjunto.

norma de derecho sustantivo; y, (iii) que ambas infracciones tengan una trascendencia significativa para la decisión¹⁸.

Cabanilla (2021, pp. 271 ss.) luego de un profundo análisis vinculado a esta causal de casación (caso 4 del art. 268), concluyó que presenta varios problemas, partiendo desde su configuración legislativa hasta la forma en que las cortes nacionales han entendido los supuestos que habilitan su procedencia. Sobre lo primero, señala que la ley ecuatoriana prohíbe textualmente la valoración probatoria en casación¹⁹, lo cual no ha impedido que, de hecho, los jueces nacionales valoren la prueba ante la verificación del error absurdo, ilógico y arbitrario. En cuanto a lo segundo, luego de un análisis extenso de criterios judiciales de todas las salas que conforman la Corte Nacional de Justicia, concluyó que existen un sinnúmero de posiciones, contradictorias, respecto al error de hecho susceptible de corrección en casación, lo cual anula toda posibilidad de previsibilidad sobre esta causal.

Estos errores que son correctamente advertidos por la autora dejan en una posición adversa al contribuyente. Piénsese que, si un tribunal declara que ha ocurrido una hipótesis de omisión, y se equivoca, el contribuyente enfrenta un sistema impreciso que pone cuesta arriba la posibilidad de impugnar la decisión, pues la legislación procesal impide que los jueces nacionales valoren prueba, pero, de hecho, cuando lo hacen, no es nada claro los supuestos en que efectivamente consideran viable la impugnación casacional. A esta dificultad se debe sumar una segunda, que es igual o más importante.

Como señala Ferrer, en efecto, si no se valora racionalmente la prueba, se violan varios derechos, por ejemplo, el debido proceso y también la presunción de inocencia. Empero, el planteamiento del autor destaca que la posibilidad de impugnar una decisión de esta naturaleza vía casación requiere de una norma jurídica sustantiva que reconozca un estándar probatorio, pues solo así sería posible evaluar si las hipótesis que han sido declaradas como probadas por el juzgador, alcanzaron el grado de suficiencia necesario para ese caso en particular. En sus palabras:

[...] la conclusión de la valoración de la prueba no puede ser más que la atribución a las distintas hipótesis en conflicto en el proceso judicial de un cierto grado de corroboración que habrá que determinar si es o no suficiente para considerar probadas las hipótesis. Ello depende de lo que establezcan las reglas jurídicas que denominamos estándares de prueba. Pues bien, *si en el proceso se declaran probados hechos para los que, teniendo en cuenta las pruebas disponibles, no se alcanza el estándar de prueba aplicable (y viceversa), el error resultante aun siendo en materia de hechos implicará la violación al derecho, en particular, de la regla que establezca el estándar de prueba aplicable.* (pp. 161-162) (énfasis añadido).

¹⁸ El art. 268.4 del COGEP prescribe que el recurso de casación procederá cuando «se haya incurrido en aplicación indebida, falta de aplicación o errónea interpretación de los preceptos jurídicos aplicables a la valoración de la prueba, siempre que hayan conducido a una equivocada aplicación o a la no aplicación de normas de derecho sustantivo en la sentencia o auto».

¹⁹ El cuarto inciso del art. 270 del COGEP prescribe: «No procede el recurso de casación cuando de manera evidente lo que se pretende es la revisión de la prueba».

En este sentido, si el único mecanismo para recurrir respecto de las decisiones dictadas contra la jurisdicción contencioso-tributaria es el recurso extraordinario de casación, y aquel necesariamente requiere un estándar de prueba para viabilizar una impugnación respecto a las inferencias probatorias y su grado de suficiencia, entonces su eficacia depende de la regla sustantiva que se fije para el efecto. Regla que hasta el momento no existe pero que es indispensable para garantizar la eficacia del recurso.

En definitiva, el escenario que enfrenta un contribuyente que es falsamente condenado por una omisión es la ausencia de remedios procesales que permitan rever y corregir la decisión. De allí la necesidad de que en un juicio tributario exista un umbral probatorio elevado para las decisiones en materia de hechos, pues si aquellas serán prácticamente inmutables, lo menos que se espera es que sean el resultado de un acervo probatorio que supere un umbral razonablemente alto. Al emplear un estándar menor, de bajas exigencias, podría ocurrir que acervos probatorios de baja calidad, insuficientes, sean el sustento de un sinnúmero de falsas condenas que en la práctica serán irreversibles.

5.5. Quinto criterio: costo social del proceso penal que resulta de una decisión adversa al contribuyente en la jurisdicción contencioso-tributaria

Finalmente, la última razón que justifica el uso del estándar probatorio planteado se extrae del coste social asociado a una imputación penal como consecuencia de una decisión que confirme la omisión en la jurisdicción contencioso-tributaria. Es justo este riesgo social el que importa para esta sección debido a los contextos latinoamericanos de discriminación, más no el riesgo de una condena en la jurisdicción penal, que es o debería ser relativamente bajo y, por ende, intrascendente para efectos de elevar o reducir el estándar probatorio.

Según el art. 298 del Código Orgánico Integral Penal (en adelante, COIP), comete el delito de defraudación tributaria quien «omita ingresos [...] para evitar el pago de los tributos debidos». La pena que se impondrá a quien sea declarado culpable de la comisión del delito de defraudación, en su modalidad de omisión, será la «pena privativa de libertad de cinco a siete años», sin embargo, si el tributo defraudado es superior a los cien salarios básicos unificados del trabajador en general, la pena podrá alcanzar los diez años de privación de libertad.

Debido a que tanto la jurisdicción contencioso-tributaria como la penal conciben a la omisión como un ilícito tributario (en sentido amplio) existe un riesgo de que, si la primera dictamina que un contribuyente incurrió en una conducta omisiva, la segunda también lo haga y, en consecuencia, exista una sanción en la esfera penal. Parece intuitivo pensar que lo uno conlleva inmediatamente a lo otro. En realidad, esto no es así.

Esta relación entre ambas jurisdicciones que aparece frecuentemente en la práctica es injustificada pues la sentencia dictada en la jurisdicción tributaria no debería

ser determinante para acreditar la hipótesis de omisión en lo penal. A lo sumo, gatilla el riesgo de imputación —pues puede y de hecho debería ser el primer insumo de una investigación penal—, pero no el de condena. Esto se explica fundamentalmente porque al discutir el delito de defraudación, se deberá probar el elemento volitivo de la conducta, es decir si existió dolo por parte del contribuyente, cuestión que para el proceso contencioso-tributario es intrascendente. Sin mencionar que en la esfera penal podría emplearse un estándar probatorio radicalmente distinto que exija otro tipo de esfuerzo probatorio que la jurisdicción tributaria no consideró.

En este orden de cosas, aunque que el riesgo de condena penal no es elevado, si lo es el de una imputación. La mera existencia del proceso penal, aunque no llegue a una condena, justifica un estándar exigente por su costo social. En Ecuador, el sistema SATJE permite acceder a la información relativa de todo proceso judicial, incluida la información sobre procesos penales. Esta libertad de acceso a la información podría generar discriminación, pues cualquiera, en cualquier momento, podría revisar si una persona natural o jurídica ha tenido conflicto con la justicia penal, independientemente del resultado del proceso (Loján-Carillo y Vázquez-Calle, 2022, pp. 748 ss.). La Corte Constitucional ha reiterado que el pasado judicial y los antecedentes penales podrían generar discriminación —es decir, son un factor de riesgo—, al punto que varias normas que exigen este tipo de información como requisitos para acceder al goce de algún derecho han sido declaradas inconstitucionales (SCC 10-14-IN/20 de 09 de junio de 2020). De allí que este criterio deba ser tomado en consideración para elevar el estándar probatorio.

Si el estándar en la justicia tributaria es elevado, las probabilidades de que exista una condena errónea son menores y, en consecuencia, se reducen también las probabilidades de que la justicia penal se active, lo cual evitaría el estigma social asociado con el pasado judicial y los antecedentes penales de una persona natural o una persona jurídica, que están latentes y a un *click* de cualquier persona con internet.

6. ¿CÓMO ADOPTAR EL ESTÁNDAR?

La decisión de adoptar un estándar probatorio es política y, por ello, debería ser el legislador quien, a través de una ley, recoja los distintos estándares de prueba en función de un debate democrático (Ferrer, 2021, pp. 149-150 y Dworkin, 1985, p. 120), sin embargo, hasta el momento, no existen señales que permitan vislumbrar algo así en el futuro cercano. Esta cuestión es ajena a los intereses del legislador, al menos por el momento. La alternativa que queda para suplir este vacío es la construcción de estándares a través de la jurisprudencia, principalmente, de la Corte Constitucional.

¿Por qué no la Corte Nacional? Pues porque las normas que los magistrados de esta judicatura deben interpretar reconocen a la concepción subjetivista de la prueba, y los estándares probatorios solo funcionan en la medida en que se adopte la concepción racionalista de la prueba. Sumado a esto, para que el recurso de casación sea,

al menos, conceptualmente efectivo para cuestionar las inferencias probatorias y los hechos probados resultantes de aquellas, se requiere la existencia de una norma que regule un estándar de prueba. Por ende, no se podría interpretar algo que aún no existe.

En este contexto, la Corte Constitucional podría adquirir un papel preponderante, pues el ámbito en que ejerce sus competencias le permite profundizar y desarrollar el contenido de los derechos constitucionalmente reconocidos en Ecuador, de un modo amplio. Prueba de ello es que en diversas situaciones que no tienen naturaleza tributaria ya ha planteado la necesidad de satisfacer un estándar probatorio para dar por probada una hipótesis fáctica.

Por ejemplo, en procesos de garantías jurisdiccionales, la Corte Constitucional señaló que debe satisfacerse el estándar probatorio de mayor probabilidad (SCC 14-19-IS/22 de 22 de junio de 2022 y 1095-20-EP/22 de 24 de agosto de 2022), con lo cual, adoptó un estándar para una jurisdicción especial, de entre las tantas existentes. Ello permite sostener que en la jurisdicción tributaria es posible incorporar un estándar de prueba para todos los casos —pero particularmente para el de la omisión—, y que la Corte lo acoja al resolver un caso concreto.

Por otro lado, la Corte Constitucional indicó que el estándar probatorio para proceder con la revocatoria del estatus de refugiado en sede administrativa requiere una alta exigencia, es decir, un estándar de más allá de toda duda razonable (SCC 215-13-EP/23 de 1 de noviembre de 2023). Aquí, se aplicó un estándar probatorio a un tipo de caso en específico, y lo interesante es que lo trasladó a sede administrativa, exigiendo que la autoridad satisfaga esta exigencia probatoria. Esto da cuenta de que también es posible pensar en estándares especiales en función del tipo de controversia tributaria que esté sobre la mesa.

En definitiva, en el peor de los escenarios en que el legislador jamás se preocupe por construir estándares probatorios, queda como alternativa la jurisprudencia. Lo importante es dejar sentada la necesidad de contar con estándares probatorios pues de no hacerlo el riesgo de la arbitrariedad seguirá latente y aquello no es admisible por los grandes y complejos valores que precautela el sistema tributario.

7. ¿CÓMO DEBERÍA APLICARSE EL ESTÁNDAR DE PRUEBA?

Para cerrar todo el análisis, es importante ejemplificar como funcionaria el estándar en casos reales. Para el efecto, se explicará cómo se utilizaría en una especie de determinación por omisión en el impuesto a la renta muy frecuente en Ecuador. Se trata de los supuestos «ingresos» por acreditaciones en cuentas bancarias²⁰. Se examinará: (i) cuál es la hipótesis de la autoridad tributaria, que deberá ser capaz de expli-

²⁰ Para revisar los fundamentos pormenorizados en estos casos, son pertinentes los documentos que obran del expediente judicial en los juicios 17510-2023-00183 y 17510-2022-00112.

car los datos disponibles y justificar las predicciones que se formulen sobre la base de la información incorporada en el expediente administrativo; y, (ii) la hipótesis alternativa del contribuyente, compatible con su ausencia de responsabilidad, que debería ser refutada solo si se aportan pruebas que le aporten algún grado de corroboración.

7.1. Hipótesis de la autoridad tributaria

Es muy frecuente que el Servicio de Rentas Internas, al hacer cruces de información con los reportes presentados por las instituciones financieras, verifique los movimientos bancarios de los contribuyentes, fundamentalmente sus depósitos en cuenta. Lo que ha ocurrido es que, sobre esa única base, se ha equiparado el concepto de «ingreso» con «acreditación en cuenta», pretendiendo gravar a todos y cada uno de esos depósitos con el impuesto a la renta.

En función de este razonamiento, la hipótesis que plantea la autoridad tributaria es que ha ocurrido una omisión en la declaración y pago del impuesto a la renta porque un contribuyente no ha considerado sus acreditaciones dinerarias en sus cuentas bancarias como ingresos susceptibles de imposición.

En función del planteamiento de Ferrer, esta hipótesis debe ser capaz de explicar los datos disponibles (los cuales deberán obrar en el expediente administrativo de la determinación y, por ende, en el expediente judicial), integrándose de forma coherente, y las predicciones de nuevos datos que la hipótesis permita formular deben haber resultado confirmadas y aportadas como pruebas al proceso.

Para acreditar la omisión del contribuyente, el SRI debe probar la ocurrencia del hecho generador del tributo. Por ello, partiendo de que la legislación ecuatoriana define que un ingreso gravado con impuesto a la renta es aquel que proviene del trabajo, capital o ambas fuentes (Ley de Régimen Tributario Interno, arts. 1 y 2), la prueba debe ser capaz de evidenciar que una acreditación dineraria percibida en una cuenta bancaria, sobre la que se pretende el ejercicio de imposición tributaria, tiene el origen antes referido.

Así, la autoridad debería desplegar todas sus facultades para, por ejemplo, verificar que un depósito determinado corresponde a una factura por honorarios, por rendimientos financieros, por pago de dividendos, o por la remuneración mensual del sujeto pasivo. Dependiendo del tipo de renta, la documentación soporte sería distinta. En esta misma línea, si se trata de un honorario profesional, seguramente tendrá que existir un contrato de prestación de servicios, un comprobante de venta y entregables del servicio. Si se trata de renta proveniente de una relación de dependencia, el contrato laboral será ese soporte.

Además, si de hecho las acreditaciones dinerarias son rentas, el SRI debería establecer cuál es el gasto en que ha incurrido el sujeto pasivo para generar este ingreso, a través de prueba. Salvo muy contadas excepciones, todo ingreso se causa por el uso de recursos del mismo sujeto pasivo para generarlos. Si esto no fuera así, no existiría

la necesidad de que las legislaciones contemplen la existencia de gastos deducibles. Esto hace que sea apenas lógico que cuando se determine el ingreso, se deba, a su vez, examinar su gasto asociado.

La autoridad tributaria también debería indagar con la contraparte en la transacción que motiva el depósito en cuenta. Por ejemplo, si se trata de un honorario profesional, quien pagó por el servicio debió registrar un gasto, y esa sería una evidencia de que la transacción que originó el depósito en cuenta es, efectivamente, un ingreso gravado con impuesto a la renta. El SRI no solo debería limitarse a indagar al sujeto pasivo determinado, sino a la contracara de sus transacciones observadas, a quienes realizan esa erogación. Con ello, se validarían las predicciones construidas a partir de la hipótesis principal.

En definitiva, para que una acreditación dineraria sea un ingreso, debe existir prueba —a cargo de la autoridad tributaria— que soporte esa conclusión en el sentido expuesto. Solo así se acreditaría la hipótesis principal, sin que ello sea suficiente para condenar al contribuyente, pues deberá refutarse su hipótesis alternativa, como se explica en el siguiente apartado.

7.2. Hipótesis del contribuyente

La defensa que se ha planteado de parte de los contribuyentes ha tenido dos vertientes: (i) una conceptual, que se ha centrado en la imposibilidad de la autoridad tributaria para presumir que toda acreditación dineraria en cuenta constituye un ingreso sujeto a imposición; y, (ii) una subsidiaria, enfocada en aportar una explicación que justifique que la acreditación dineraria no constituye un ingreso gravado con impuesto a la renta.

Desde un enfoque de la segunda línea de defensa, la hipótesis alternativa formulada por la defensa de los contribuyentes, que debe ser necesariamente refutada siempre que exista alguna evidencia que le aporte corroboración, es que en la declaración del impuesto a la renta —de haberla— se registró y cuantificó total y adecuadamente el hecho imponible del impuesto a la renta, sin que la acreditación dineraria altere en ningún sentido este escenario.

Entre las muchas justificaciones que se han esgrimido en determinaciones de este tipo, los contribuyentes han sostenido que gran parte de las acreditaciones dinerarias corresponden a rentas exentas, pagos de capital de préstamos y transacciones entre cuentas del mismo sujeto pasivo. Por tal razón, esas acreditaciones en cuenta jamás podrían tener la categoría de un ingreso susceptible de gravamen tributario. Como soporte, se han presentado contratos de mutuo, escrituras de compraventa de inmuebles, y explicaciones detalladas de las transacciones.

Para que prevalezca la hipótesis del SRI, debe refutarse esta hipótesis alternativa que es compatible con la ausencia de responsabilidad del contribuyente, siempre que encuentre corroboración en evidencia del tipo que fuere, pues demuestra que el

tributo fue correctamente cuantificado y consideró todas las rentas gravadas, excluyendo a aquellas acreditaciones dinerarias que no calzan en el concepto legal de renta. Es importante insistir en que esta hipótesis, favorable al contribuyente, no debe ser puesta bajo la lupa del mismo estándar exigido para el fisco; debe ser necesariamente menos exigente por tratarse de la inocencia o ausencia de responsabilidad del sujeto pasivo.

Por lo expuesto, la autoridad tributaria no podría aducir que su hipótesis de omisión se cumplió, sin refutar motivadamente esta segunda junto con toda la evidencia que acredite el cumplimiento tributario del contribuyente.

7.3. Conclusión

Si no se supera el estándar planteado, la autoridad tributaria, sobre quien recae la carga objetiva de la prueba, tendría que perder el proceso por no alcanzar el grado de suficiencia probatoria exigido para acreditar la ocurrencia de la omisión.

Como se desarrolló en este artículo, el estándar 3 tiene lógica y sentido porque es el Estado quien, en ejercicio de su facultad determinadora, debe acreditar la ocurrencia del hecho imponible para luego cuantificarlo y perseguir su cobro. Bajo esta consigna, y por los poderes que ostenta el fisco, es razonable exigir una investigación exhaustiva acompañada de un cúmulo cuantioso de evidencia como resultado.

Al reducir el estándar, podría generarse un incentivo negativo que incite a que la facultad determinadora no se ejerza con diligencia y, por ende, el fisco exija el pago de impuestos no debidos sin la evidencia necesaria para que se desprenda el poder tributario sobre los ciudadanos. Es decir, se legitimaría la confiscación, sin que exista un remedio procesal útil que permita corregir la decisión y que, además, vendrá acompañada de un coste social asociado a una posible imputación penal.

8. BIBLIOGRAFÍA

- Aguirre, A. Arrobo, S. Ruiz, S. Puertas, P. (2022). La inseguridad jurídica por la falta de la doble instancia en el proceso contencioso tributario en el Ecuador. Universidad Nacional del Altiplano. *Revista de Derecho*, 7(2), 3-18. <https://www.redalyc.org/journal/6718/671871944002/html/>
- Barcena, A. (2012). *La causalidad en el derecho de daños* (Capítulo IV La causalidad por omisión, p. 181-240) [Tesis Doctoral, Universitat de Girona]. <http://hdl.handle.net/10803/108448>
- Cabanilla, M. (2021). *Valoración probatoria en el recurso extraordinario de casación ecuatoriano. Análisis doctrinario, normativo y jurisprudencial*. Corporación de Estudios y Publicaciones.
- Centro Interamericano de Administraciones Tributarias. Cooperación Alemana. Banco Interamericano de Desarrollo. (2015). *Modelo de Código Tributario del CIAT: un enfoque basado en la experiencia iberoamericana*. Recuperado de https://www.ciat.org/Biblioteca/DocumentosTecnicos/Espanol/2015_Modelo_Codigo_Tributario_CIAT.pdf
- Danós, J. (2000). El proceso contencioso administrativo en materia tributaria. *THEMIS Revista De Derecho*, 41, 73-95. <https://revistas.pucp.edu.pe/index.php/themis/article/view/11653>

- Dei Vecchi, D. (2019). El carácter presuntivo de las presunciones absolutas. *Revus, Journal for Constitutional Theory and Philosophy of Law*, 38, 1-33. <http://journals.openedition.org/revus/5333>
- Diccionario Panhispánico del Español Jurídico. (2024). *Omisión*. <https://dpej.rae.es/lema/omisi%C3%B3n>
- Ferrer Beltrán, J. (2018). Los hechos en la casación penal. *Revista pensamiento penal*, 48, 153-175. <https://www.pensamientopenal.com.ar/doctrinal/46775-hechos-casacion-penal>
- Ferrer Beltrán, J. (2021). *Prueba sin convicción. Estándares de prueba y debido proceso*. Marcial Pons.
- Folco, C. (2011). *Procedimiento tributario, naturaleza y estructura* (Tercera edición ampliada y actualizada). Rubinal-Culzoni Editores.
- Garcés, R. (1980). *La caducidad en el Sistema Tributario Ecuatoriano*. Ediciones de la Universidad Católica.
- Guerrero, F. (2019). *Nueva visión del Derecho Administrativo: sustentada en el Código orgánico Administrativo y normativa conexas*. Editora Jurídica.
- Giuliani Fonrouge, C. (1997). *Derecho Financiero* (volumen I). Depalma.
- González, J. (2000). Facultades de la administración tributaria en materia de determinación de tributos. *THEMIS Revista de Derecho*, 41, 17-28. <https://revistas.pucp.edu.pe/index.php/themis/article/view/11648>
- Lascano, G. (2017). El Proceso Lógico - Jurídico de la Convicción del Juez Tributario. *Revista Facultad De Jurisprudencia*, 1(1). <https://doi.org/10.26807/rjf.v1i1.10>
- Letelier, R. (2018). El precio del *statu quo*. Sobre el estándar probatorio en las sanciones administrativas. *Revista de Derecho (Valdivia)*, 31(1), 209-229.
- Loján, S. y Vázquez, J. (2022). La discriminación y el pasado judicial. Estudio del Sistema Informático de Trámites Judiciales SATJE. *Revista arbitrada interdisciplinaria KOINONIA*, 7(2), 737-754.
- Mejía, Al. (2013). La prueba en el proceso tributario ecuatoriano. En M. Albán Zambonino, *Tópicos contemporáneos del Derecho Tributario. Homenaje al Instituto Ecuatoriano de Derecho Tributario en sus 50 años* (pp. 57-78). Cevallos Editora Jurídica.
- Narváez, M. (2010). Causalidad y omisión en la actuación médica. Rompecabezas conceptual para el enjuiciamiento penal. *Diritto & questione pubbliche*, 10, 225-226.
- Neira, E. (2016). *La jurisdicción contencioso-administrativa: reflexión sobre sus disfuncionalidades*. Editorial USFQ.
- Pérez, E. (2021). *El acto administrativo en el Ecuador*. Corporación de Estudios y Publicaciones.
- Pino, D. (2023). *Manual de derecho tributario formal*. Ediciones Legales EDLE S.A.
- Piza, J. (2010). *Curso de Derecho Tributario, procedimiento y régimen sancionatorio* (1.ª edición). Universidad Externado de Colombia.
- Reyes, J. y Vallejo, A. (2016). Ecuador. En S. Whitehead (ed.), *The tax disputes and litigation review* (4th edition, pp. 110-120). Law Business Research.
- Rezzoagli, B. (2009). Ilícitos tributarios Diferenciación entre evasión, defraudación y elusión fiscal. *Amicus Curiae: revista electrónica de la Facultad de Derecho*, 2(2). <https://www.revistas.unam.mx/index.php/amicus/article/view/13543>
- Ríos, G. (2005). El ilícito tributario. En M. de J. Alvarado Esquivel (coord.), *Manual de Derecho Tributario*. Ed. Porrúa.
- Rodríguez, J. (2012). *Prescripción extintiva y caducidad en el derecho civil. Estudio de su régimen jurídico en el Código Civil y en la Ley de Relaciones de Consumo* (3.ª ed.). Facultad de Cultura Universitaria.
- Rodríguez, M. (2007). *La prueba en el derecho tributario. Formación y preconstitución de la prueba en el derecho tributario* (1.ª ed.). Editorial Aranzadi, S.A.
- Rosen-Zvi, I. y Fisher, T. (2008). Overcoming procedural boundaries. *Virginia Law Review*, 94(1), 79-155. <http://www.jstor.com/stable/25470553>
- Saccone, M. (2002). *Manual de Derecho Tributario*. Editorial La ley.
- Simone, C. (2017). El proceso contencioso tributario y el control de legalidad de la actuación administrativa. En E. M. de la Guerra Zúñiga (dir.), *Derecho financiero y derecho tributario: ensayos en honor del profesor José Vicente Troya Jaramillo* (volumen 1, pp. 449-467). Instituto Ecuatoriano de Derecho Tributario.
- Troya, J. (2014). *Manual de Derecho Tributario* (1.ª ed.). Corporación de Estudios y Publicaciones.

- Tuzet, G. (2013). *The omission dilemma*. Bocconi Legal Studies Research Paper, 2887393. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2887393
- Valdés Costa, R. (2016). *Curso de Derecho Tributario* (4.ª ed.). Editorial B de F Montevideo - Buenos Aires.
- Van Hasselt, R. (2015). El estándar de prueba en materia infraccional. *Revista De Estudios Tributarios*, 13, 209–236. <https://revistaestudiostributarios.uchile.cl/index.php/RET/article/view/39875>
- Vidal, J. (1985). *Derecho Administrativo*. Librería Temis Editorial.

Leyes y jurisprudencia

- Corte Nacional de Justicia (Ecuador). *Resolución de 4 de mayo de 2011. Registro oficial no. 471 de 16 de junio de 2011*. <https://www.registroficial.gob.ec/index.php/registro-oficial-web/publicaciones/suplementos/item/3458-suplemento-al-registro-oficial-no-471>
- Corte Suprema (Argentina) *Fallo 186-170*
- Ecuador (2004). *Ley de Régimen Tributario Interno [LRTI]*. <https://www.ces.gob.ec/lotaip/2018/Agosto/Anexos-literal-a2/LEY%20DE%20REGIMEN%20TRIBUTARIO%20INTERNO,%20LRTI.pdf>
- Ecuador (2005). *Código de procedimiento civil*. <https://www.planificacion.gob.ec/wp-content/uploads/2015/04/Codigo-de-Procedimiento-Civil.pdf>
- Ecuador (2005). *Código Tributario*. https://www.sri.gob.ec/sri-portlet-biblioteca-alfresco-internet/descargar/00682c30-727a-4268-b095-f485d226ce49/Codigo_Tributario_10082023.pdf
- Ecuador (2015). *Código general de procesos [COGEP]*. <https://www.telecomunicaciones.gob.ec/wp-content/uploads/2018/09/Codigo-Org%20nico-General-de-Procesos.pdf>
- Ecuador (2014). *Código orgánico integral penal [COIP]*. https://www.defensa.gob.ec/wp-content/uploads/downloads/2021/03/COIP_act_feb-2021.pdf
- Sentencia de la Corte Constitucional (Colombia) C-293/20, de 5 de agosto de 2020. [hbhttps://www.corteconstitucional.gov.co/relatoria/2020/C-293-20.htm](https://www.corteconstitucional.gov.co/relatoria/2020/C-293-20.htm).
- Sentencia de la Corte Constitucional (Ecuador) 10-14-IN/2,0 de 09 de junio de 2020.
- Sentencia de la Corte Constitucional (Ecuador) 47-15-IN/21, 10 de marzo de 2021.
- Sentencia de la Corte Constitucional (Ecuador) 14-19-IS/22, de 22 de junio de 2022.
- Sentencia de la Corte Constitucional (Ecuador) 1095-20-EP/22, de 24 de agosto de 2022.
- Sentencia de la Corte Constitucional (Ecuador) 46-18-IN/23, de 6 de septiembre de 2023.
- Sentencia de la Corte Constitucional (Ecuador) 215-13-EP/23, de 1 de noviembre de 2023.
- Sentencia de la Corte Constitucional (Ecuador) 10-21-IN/23, de 22 de noviembre de 2023.
- Sentencia de la Corte Constitucional (Perú) 0041-2004-AI/TC, del 11 de noviembre de 2004. <https://www.tc.gob.pe/jurisprudencia/2005/00041-2004-AI.pdf>
- Sentencia de la Corte Constitucional (Perú) 00053-2004-PI/TC, de 16 de mayo de 2005. <https://tc.gob.pe/jurisprudencia/2005/00053-2004-AI.pdf>
- Sentencia de la Corte Nacional de Justicia (Ecuador) 0119-2010 (Sala de lo Contencioso Administrativo), de 19 de abril de 2010.
- Sentencia de la Corte Nacional de Justicia (Ecuador) 0339-2010 (Sala de lo Contencioso Administrativo), de 1 de octubre de 2010.
- Sentencia de la Corte Nacional de Justicia (Ecuador) 01803-2019-00258 (Sala de lo Contencioso Administrativo), de 21 de septiembre de 2020.
- Sentencia de la Corte Nacional de Justicia (Ecuador) 17811-2013-0648 (Sala de lo Contencioso Administrativo), de 10 de febrero de 2022.
- Sentencia de la Corte Nacional de Justicia (Ecuador) 09802-2020-00424 (Sala de lo Contencioso Administrativo), de 19 de septiembre de 2022.
- Sentencia de la Corte Nacional de Justicia (Ecuador) resolución 15-2009 (Sala de lo Contencioso Tributario), de 18 de febrero de 2009.
- Sentencia de la Corte Nacional de Justicia (Ecuador) 17510-2017-00528 (Sala de lo Contencioso Tributario), de 11 de junio de 2016.

- Sentencia de la Corte Nacional de Justicia (Ecuador) 09501-2017-00365 (Sala de lo Contencioso Tributario), de 29 de junio de 2018.
- Sentencia de la Corte Nacional de Justicia (Ecuador) 13802-2017-00131 (Sala de lo Contencioso Tributario), de 27 de noviembre de 2018.
- Sentencia de la Corte Nacional de Justicia (Ecuador) 17510-2017-00427 (Sala de lo Contencioso Tributario), de 6 de mayo de 2019.
- Sentencia de la Corte Nacional de Justicia (Ecuador) 09501-2018-00175 (Sala de lo Contencioso Tributario), de 20 de mayo de 2021.
- Sentencia de la Corte Nacional de Justicia (Ecuador) 09501-2018-00057 (Sala de lo Contencioso Tributario), de 31 de mayo de 2021.
- Sentencia de la Corte Nacional de Justicia (Ecuador) 09501-2017-00791 (Sala de lo Contencioso Tributario), de 21 de junio de 2021.
- Sentencia de la Corte Nacional de Justicia (Ecuador) 09501-2018-00281 (Sala de lo Contencioso Tributario), de 30 de junio de 2021.
- Sentencia de la Corte Nacional de Justicia (Ecuador) 09501-2017-00736 (Sala de lo Contencioso Tributario), de 22 de julio de 2021.
- Sentencia de la Corte Nacional de Justicia (Ecuador) 17510201900004 (Sala de lo Contencioso Tributario), de 20 de agosto de 2021 .
- Sentencia de la Corte Nacional de Justicia (Ecuador) 09501-2018-00126 (Sala de lo Contencioso Tributario), de 20 de septiembre de 2021 .
- Sentencia de la Corte Nacional de Justicia (Ecuador) 175102019-00297 (Sala de lo Contencioso Tributario), de 20 de mayo de 2022.
- Sentencia de la Corte Nacional de Justicia (Ecuador) 17510-2018-00167 (Sala de lo Contencioso Tributario), de 29 de septiembre de 2022.
- Sentencia de la Corte Nacional de Justicia (Ecuador) 17510-2019-00389 (Sala de lo Contencioso Tributario), de 6 de octubre de 2022.
- Sentencia de la Corte Nacional de Justicia (Ecuador) 17510-2018-00418 (Sala de lo Contencioso Tributario), de 25 de octubre de 2022.
- Sentencia de la Corte Nacional de Justicia (Ecuador) 17510-2018-00383 (Sala de lo Contencioso Tributario), de 17 de febrero de 2023.
- Sentencia de la Corte Nacional de Justicia (Ecuador) 17510-2016-00192 (Sala de lo Contencioso Tributario), de 27 de abril de 2023.
- Sentencia de la Corte Nacional de Justicia (Ecuador) 17510-2019-00311 (Sala de lo Contencioso Tributario), de 2 de junio de 2023.
- Sentencia de la Corte Nacional de Justicia (Ecuador) 17510-2020-00285 (Sala de lo Contencioso Tributario), de 14 de noviembre de 2023.
- Sentencia de la Corte Nacional de Justicia (Ecuador) recurso no. 383-2011 (Sala de lo Contencioso Tributario), de 31 de octubre de 2013.
- Sentencia de los Tribunales Distritales de lo Contencioso-Tributario (Ecuador), juicio 17504-2011-011, de 25 de junio de 2020.
- Sentencia de los Tribunales Distritales de lo Contencioso-Tributario (Ecuador), juicio 17510-2023-00028, de 23 de noviembre de 2023.
- Sentencia de los Tribunales Distritales de lo Contencioso-Tributario (Ecuador), juicio 17510-2023-00028, de 27 de noviembre de 2023.
- Sentencia de los Tribunales Distritales de lo Contencioso-Tributario (Ecuador), juicio 17510-2022-00259, de 13 de diciembre de 2023.
- Tribunal Distrital de Lo Contencioso Tributario (Ecuador), juicio 17510-2023-00183.
- Tribunal Distrital de Lo Contencioso Tributario (Ecuador), juicio 17510-2022-00112.

LA PRUEBA PERICIAL CIENTÍFICA EN MANCHAS DE SANGRE: UNA REFLEXIÓN SOBRE SU APLICACIÓN EN EL PROCESO ESPAÑOL. SESGOS Y RECOMENDACIONES

María Hernández Moreno

Universidad Isabel I

mahemo94@hotmail.com

RESUMEN: La prueba pericial científica, trascendental para la toma de decisiones, se caracteriza por ser rigurosa y objetiva y por recoger información de interés sobre cuestiones ajenas al conocimiento jurídico. A pesar de su relevancia procesal, no todas las disciplinas cuentan con el mismo desarrollo, pudiendo producirse errores en la elaboración y valoración de informes de aquellas menos demandadas y, por ende, menos respaldadas por la doctrina, como ocurre con el análisis de patrones de manchas de sangre, apenas expandido en la práctica a pesar de su enorme valía para la reconstrucción de los hechos. Por medio de este artículo se estudiará y presentará su situación en España, analizándose jurisprudencia y normativa al respecto tanto de su contenido como de la formación requerida a sus expertos. Con ello, podrán aportarse recomendaciones fundamentadas que faciliten el desarrollo de estas pericias de forma garante, al minimizar los obstáculos y las limitaciones observadas.

PALABRAS CLAVE: ciencias forenses; manchas de sangre; proceso judicial; prueba científica; prueba pericial.

SCIENTIFIC EXPERT REPORT IN BLOODSTAINS: A REFLECTION ON ITS APPLICATION IN THE SPANISH COURT. BIASES AND RECOMMENDATIONS

ABSTRACT: Scientific expert report, which is essential for decision-making, is characterized by being rigorous and objective and by being able to collect information of interest on issues outside of legal knowledge. Despite its procedural relevance, not all disciplines have the same development, and errors may occur in the preparation and evaluation of reports of those that are less in demand and, therefore, less supported by doctrine. It occurs with bloodstain pattern analysis; hardly expanded in practice even though its enormous value for the reconstruction of the facts. Through this article, its situation in Spain will be studied and presented, analyzing jurisprudence and regulations regarding both its content and the training required for its experts. With this, well-founded recommendations can be provided that facilitate the development of these expert reports in a guaranteed manner, by minimizing the obstacles and limitations observed.

KEYWORDS: bloodstains; expert report; forensic sciences; judicial process; scientific report.

SUMARIO: 1. INTRODUCCIÓN.— 2. ANÁLISIS DE PATRONES DE MANCHAS DE SANGRE: PRESENTACIÓN DE LA DISCIPLINA Y DE SU DESARROLLO EN ESPAÑA.— 3. LA PRUEBA PERICIAL CIENTÍFICA Y SU VALORACIÓN.— 4. ESTADO DEL INFORME PERICIAL EN MANCHAS DE SANGRE EN ESPAÑA: 4.1. Fundamentos de la pericial en manchas de sangre; 4.2. Contenido del informe pericial; 4.3. Revisión jurisprudencial.— 5. REFLEXIONES Y RECOMENDACIONES DE MEJORA.— BIBLIOGRAFÍA.

1. INTRODUCCIÓN

Uno de los grandes objetivos de la investigación criminal descansa en la reconstrucción de los hechos, facilitando con ello su resolución al poder determinar qué ocurrió, cómo, dónde, cuándo y quién o quiénes se vieron implicados. Para lograrlo, y atendiendo al carácter científico de las disciplinas encargadas del estudio de esos elementos de interés, se acude a actividades periciales, especialmente valiosas para probar ciertos hechos que, de otro modo, difícilmente podrían demostrarse, convirtiéndose en una de las pruebas cada vez más demandadas en los procesos judiciales (Meza, 2024).

Cuando se investigan hechos que atentaron directamente contra la vida o la integridad física, en los que la sangre se convierte en el vestigio protagonista, resulta de gran interés atender a los resultados recogidos en el informe pericial centrado en la reconstrucción de los hechos y respaldado desde el análisis de patrones de manchas de sangre. Disciplina que determinará, a partir de la disposición de los depósitos y los restos hemáticos presentes en el lugar de interés, la manera en la que se desarrolló el episodio sangriento en cuestión, aportando valor a la investigación al lograr concretar posiciones de víctima y victimarios, posibles manipulaciones en el escenario o sus elementos, líneas temporales descritas, afirmaciones o contradicciones en los testimonios de los involucrados, alteraciones y sustracciones de objetos, etc., contribuyendo enormemente a la reproducción de lo ocurrido, y por ende, a la resolución efectiva de cada investigación criminal.

Sin embargo, y a pesar de todo lo anterior, apenas se encuentra desarrollada como materia pericial en nuestro país, precisándose de un análisis en profundidad, tanto de su contenido como de su fiabilidad y rigor, para poder así identificar su valía y demostrar que cumple con los requisitos propios del informe como medio de prueba y que, por tanto, podrá ser valorado por el órgano competente de forma garante.

El presente artículo estudiará, atendiendo a todo lo anterior, el estado actual de la pericia en manchas de sangre en el proceso español, aportando información de relevancia sobre su contenido y sobre la rigurosidad de sus resultados, valorando, en definitiva, su validez probatoria. Para ello se abordarán, en primer lugar, unas consideraciones generales sobre la disciplina y el concepto de prueba científica, analizando el respaldo normativo y jurisprudencial del informe pericial y de sus características y contenido en relación con la disciplina de interés, finalizando el artículo con una serie de sugerencias de mejora que podrían implementarse en aras de garantizar en todo momento su rigor y calidad y que se centrarán, en gran medida, en apostar por un modelo educativo (Cando Shevchukova, 2023), basado en enfatizar la necesidad de dotar de información, estándares y formación homogénea a los especialistas por un lado, y en profundizar la investigación de las pruebas para que las decisiones presentadas en estos documentos sean lo más racionales y correctas posibles (Gascón, 2020, p. 250-261; Vázquez, 2015, p. 149 y ss.).

2. ANÁLISIS DE PATRONES DE MANCHAS DE SANGRE: PRESENTACIÓN DE LA DISCIPLINA Y DE SU DESARROLLO EN ESPAÑA

En primer lugar, y antes de entrar a valorar la trascendencia del informe pericial en manchas de sangre, conviene concretar los objetivos de la disciplina y poner en valor las aportaciones que puede hacer al proceso judicial, logrando así comprender también sus posibles limitaciones y los márgenes normativos en los que, en algunas ocasiones, se encuentra.

Para ello es necesario exponer que el análisis de patrones de manchas de sangre forma parte de las denominadas ciencias forenses¹ al centrarse en el estudio y la interpretación de uno de los vestigios más predominantes en los escenarios criminales, la sangre, con el fin de auxiliar a la administración de justicia en el proceso de toma de decisiones. Su objetivo fundamental se asienta, por tanto, en lograr la reconstrucción del episodio sangriento a través de la interpretación de la morfología de las manchas y los patrones que puedan observarse en el lugar de interés con la intención

¹ Conclusión evidente atendiendo a lo recogido por De Barros *et al.* (2021) que concretan que serán aquellas que empleen conocimientos científicos y técnicas para investigar delitos y otros asuntos legales con el objetivo de resolver problemas de seguridad pública. Así, a partir del estudio de las evidencias recogidas en el ámbito de la investigación, estas ciencias ayudarán a dilucidar un determinado delito, creando hipótesis sobre lo ocurrido.

de aportar información de valor sobre qué ocurrió, cómo, dónde y quién o quiénes se vieron involucrados. No será relevante, en este caso, atender al contenido ni a la información genética que encierre; únicamente se valorará la disposición de cada mácula, su tamaño, forma, regularidad, etc., o lo que es lo mismo, el continente del vestigio y el aspecto que haya adoptado en cada caso.

Su desarrollo en los últimos años², marcado por grandes avances y mejoras en las técnicas y la tecnología empleadas para su correcta aplicación en los diferentes casos criminales, han convertido a esta disciplina en una de las más relevantes en países como Estados Unidos, Francia, Reino Unido o Países Bajos (Hernández Moreno, 2022). En el caso de España, y como irá exponiéndose a lo largo del artículo, todavía queda un largo camino por recorrer para que su reconocimiento esté al nivel de su relevancia forense.

Ese desigual crecimiento entre países, unido a su desarrollo excesivamente reciente, pues no ha sido hasta el inicio del siglo XXI cuando ha gozado del mayor impulso de toda su historia (Illes, 2001), se agrava, en nuestro caso, por una serie de cuestiones que pueden reducirse a la ausencia de formación y normativa específica y al desconocimiento generalizado de la existencia y las funciones y aplicaciones de la propia disciplina. Y es que, más allá de una Decisión Marco (2009/905/JAI) del Consejo de la Unión Europea de 30 de noviembre de 2009, que aborda la necesidad de introducir acreditaciones para aquellos que presten servicios forenses por trabajar en laboratorios con información tan delicada como ADN y datos dactiloscópicos, no se cuenta con más legislación que pueda ser de aplicación a esta disciplina en nuestro país. Tampoco hay estándares de actuación a los que deban acudir los peritos, ni sistemas que determinen cuál debe ser su grado de especialización o qué nivel de estudios deben haber completado para acreditar que realmente son expertos en esta materia y que tienen el nivel suficiente para desarrollar un peritaje sobre manchas de sangre en la escena; normas que sí existen en otros países europeos como Francia o Países Bajos y, por su puesto, en EEUU (Hernández Moreno, 2022).

Esto se traduce en una carencia de profesionales acreditados, al contarse con un número reducido de ellos entre las diferentes Fuerzas y Cuerpos de Seguridad (en adelante FFCCSS)³. Problemática que se agrava si se atiende a la desigualdad reinante, como se comentaba anteriormente, en la formación recibida y en los niveles de especialización en cada caso⁴.

² Si bien los primeros documentos sobre la materia empezaron a publicarse en Europa en la década de 1800, no es reconocida como disciplina forense hasta 1939, despuntando en 1971 gracias a Hebert Leon MacDonell, considerado el padre del análisis de patrones de manchas de sangre (Bevel y Gardner, 2008; James *et al.*, 2005; James y Eckert, 1999).

³ Siguiendo la recomendación de la Asociación Internacional de Patrones de Manchas de Sangre (IABPA por sus siglas en inglés) de haber cursado al menos 40h de forma homologada a través de la realización de su curso básico en alguno de los centros adscritos y reconocidos por la asociación.

⁴ En función de los profesionales europeos que ofrezcan la formación, se cuenta con un itinerario formativo diferente una vez se finaliza el curso básico de 40h.

A ello debe sumarse el escaso desarrollo de la disciplina en nuestro territorio y, por tanto, de la experiencia de sus miembros. Es importante entender que, aunque en los últimos años haya crecido su protagonismo, quedando plenamente integrada en algunas unidades de FFCCSS relacionadas, por ejemplo, con la Inspección Técnico Ocular (ITO) del escenario criminal, todavía no está consolidada como una disciplina forense autónoma y, por tanto, no goza del papel que ocupan otras como la genética forense o las trazas instrumentales, claramente afianzadas en las investigaciones criminales. Obstáculo que recogía Moreno Lopera en el único estudio realizado hasta la fecha sobre estas cuestiones y en el que afirmar que «el 100 por 100 de los profesionales (encuestados) no ha realizado ningún tipo de actividad práctica sobre la interpretación de los patrones de manchas de sangre», materializándose ese desconocimiento de la materia igualmente en el desarrollo de informes periciales pues «más del 90 por 100 no ha declarado en un juicio [...] ni conoce a algún compañero que haya declarado como perito». Lo mismo ocurriría a la hora de aplicar los conocimientos obtenidos a los casos reales que surjan de forma cotidiana, de hecho «más de un 85 por 100 asegura que su unidad no ha utilizado la interpretación de manchas de sangre en la investigación de homicidios, ni han intervenido en casos donde se estudiaran dichas manchas» (Moreno Lopera, 2015).

Es cierto que desde la publicación de ese estudio en el año 2015 hasta la actualidad se han realizado numerosos cursos formativos, tanto en cuerpos nacionales como autonómicos y que, por tanto, ha de asumirse que, en caso de repetirse el estudio, las cifras recogidas serían mucho más alentadoras. Sin embargo, ello no supone que los datos anteriores no tengan valor; han de tomarse como base para poder comprender de qué punto se parte y qué elementos pueden seguir lastrando su desarrollo a día de hoy, pues, recordemos, su implementación no está generalizada ni puede ser tratada como disciplina independiente.

3. LA PRUEBA PERICIAL CIENTÍFICA Y SU VALORACIÓN

Una vez introducida la disciplina objeto de estudio y contextualizada su situación en nuestro país, conviene profundizar en su carácter científico para comprender su valía procesal y la forma en la que quedará valorada por el órgano competente.

Si atendemos a los diferentes elementos presentes en cualquier escenario criminal, y especialmente en aquellos en los que se haya atentado contra la vida o la integridad física, de forma general, y más allá de su variedad en cuanto a tipologías y características, puede afirmarse que presentan un punto en común nada desdeñable: han de ser inspeccionados y estudiados desde disciplinas científicas, otorgándoles la condición de prueba pericial científica una vez forman parte del proceso judicial. En estas y otras ocasiones, los hechos objeto de prueba pueden llegar a requerir, por su naturaleza, de conocimientos de tipo científico, artístico, técnico o práctico para su correcta interpretación (Abel Lluch y Picó i Junoy, 2009), evidenciando con ello la finalidad del informe pericial: suplir la falta de conocimientos no jurídicos del juez. Esto se tra-

duce en una necesidad evidente de precisar de personal especializado en cada materia de interés para que pueda auxiliar, de forma correcta, al tribunal competente durante el proceso de toma de decisiones.

En el caso de las evidencias del escenario criminal, por norma general, los conocimientos que deberán haberse adquirido para su correcta interpretación serán de tipo científico, es decir, versarán sobre ciencia, entendida como un conjunto de conocimientos razonados que versan sobre objetos de una misma naturaleza, han sido obtenidos mediante la aplicación sistematizada de métodos concretos y pueden ser probados (Egg, 1977).

Otras definiciones comprenden la ciencia como un método experimental o un instrumento de rigor que permite investigar fenómenos sociales al formular hipótesis que serán refutadas o verificadas a través de la experimentación, pero también como un acumulado de conocimientos que se han adquirido a lo largo de la investigación propiamente dicha (Sánchez Rubio, 2016). Es decir, entienden la ciencia como una forma de actividad social que pretende producir, distribuir y aplicar diferentes conocimientos al presentarse como una institución social al servicio de las necesidades y posibilidades de cada sociedad (Asencio Cabot, 2014), yendo más allá de un mero conjunto de conceptos o teorías. Será, además, sistemática, acumulativa, metódica, comprobable, especializada, abierta y el resultado de una investigación concreta (Feynman, 2001).

En definitiva, el concepto de ciencia se caracteriza por encerrar veracidad e información objetiva, al entenderse que, en todo momento, cualquiera que sea la exploración que se desarrolle, está respaldada por una metodología reproducible que garantiza, al poder ser contrastada y repetida por terceros, que no ha habido manipulaciones ni vicios en el desarrollo de dicha investigación ni en sus resultados. Ese método científico otorga, además, fiabilidad a cada estudio, entendiéndose como un estándar por el cual pueden demostrarse y evaluarse hipótesis científicas, al tiempo que permite estudiar y averiguar cómo funciona una determina cosa (Vázquez, 2014), y se considera, a pesar de los desacuerdos filosóficos y científicos (Andersen y Hepburn, 2015) como «verdadero, objetivo y neutral» (Cippitani *et al.*, 2021), si bien, y como veremos más adelante, la labor de los especialistas en ciencias forenses no está exenta de sesgos.

Aplicando toda esa información al ámbito que nos concierne, el de las ciencias forenses (al que pertenece el análisis de patrones de manchas de sangre), podrá establecerse que la prueba será científica en aquellos casos en los que el proceso para su obtención exija de una experiencia concreta que permita alcanzar resultados y conclusiones, al implantar un cierto método para el que habrá sido necesario acometer una serie de estudios o prácticas previas y concretas sobre dicha materia (Gonaíni, 2012).

Es decir, una prueba adquiere el valor de científica en el momento en el que, para poder estudiar la cosa o el hecho controvertido objeto de dicha pericial, sea necesario

acudir a uno o varios métodos de actuación científica respaldado por experiencias y estudios concretos sobre una materia. Se trataría, por tanto, de una variante de la prueba pericial común, que se caracteriza por la especialidad del experto encargado de su elaboración, al desarrollarla en base a los aportes y métodos de una materia científica concreta, acogiéndose a sus normas y avances tecnológicos que serán aplicados en el examen de la cosa o el hecho objeto de estudio. En el caso del análisis de manchas y patrones de sangre, y como quedará descrito más adelante, los expertos deberán acudir a diversas técnicas y metodologías propias de la materia para poder desgranar la información que encierran esos depósitos hemáticos e incluso, para poder concretar su localización en ciertos escenarios, precisándose de una formación especializada que les habilite para que puedan proceder, de forma adecuada y certera, a la correcta interpretación *in situ* y al posterior estudio de esos vestigios.

Estas pruebas científicas, además, tienen una serie de características particulares: son pruebas de colaboración, pues normalmente, para poder llevarlas a cabo, se hace necesaria la actuación y cooperación de terceros; sus fuentes son muy dinámicas y cambian de forma constante, al tiempo que avanza la tecnología de su campo de estudio; en muchas ocasiones, para probar la autenticidad del elemento objeto de la pericia en sí mismo, se hace necesario acudir a otra prueba complementaria (Peyrano, 2015). Desde el análisis de patrones de manchas de sangre y otras tantas disciplinas científicas pertenecientes a las ciencias forenses, destaca la paridad en esos puntos comunes: necesitan la colaboración con investigadores de otras disciplinas para concluir con éxito cada investigación (fotografía forense; genética forense; etc.); el avance en las técnicas, cada vez más tecnológicas, obligan a los analistas a estar en formación continua; en ciertas ocasiones, además, los estudios deben quedar complementados con otras pruebas que garanticen la objetividad del proceso (pruebas de ADN para demostrar que las manchas analizadas o localizadas son realmente sangre humana y pertenecen a un cierto individuo).

Sin embargo, y pesar de todo ello, el valor de estas pericias no se da por sentado, pues dependerá de la validez científica del método que se haya empleado y de si se han implementado las tecnologías y los controles de calidad pertinentes y apropiados en cada ocasión (Moreno Lopera, 2015). Puede afirmarse, por tanto, que, si bien las pruebas periciales científicas permiten acceder a la verdad sobre unos hechos concretos en la gran mayoría de las ocasiones (Vargas Ávila, 2010), su valor no es innato, y deberán cumplirse las normas y métodos científicos de cada disciplina para demostrar los resultados que se recojan en el informe de una manera objetiva, eficaz y, sobre todo, veraz, prestando con ello auxilio en la toma de decisiones del proceso de manera adecuada (Moreno Lopera, 2015). Y es que, tal y como recoge la LECrim en su artículo 348: «El tribunal valorará los dictámenes periciales según las reglas de la sana crítica», concretando, en base al artículo 741 de esa misma ley que: «El Tribunal, apreciando, según su conciencia las pruebas practicadas en el juicio, las razones expuestas por la acusación y la defensa y lo manifestado por los mismos procesados, dictará sentencia dentro del término fijado en esta Ley».

Es decir, que el juez goza de la libertad necesaria como para hacer una valoración del informe atendiendo al resto de las pruebas y haciendo una apreciación de todas ellas en conjunto, prestando atención, además, a dos aspectos primordiales respaldados por la jurisprudencia (Abel Lluch y Picó i Junoy, 2009; Cameriere *et al.*, 2013; Meneses Pacheco, 2008):

— Los razonamientos que contengan los dictámenes periciales, así como a los que se hubieran vertido por parte del perito durante el plenario, deberán ser atendidos por el juez, pudiendo rechazar el resultado de uno de esos informes o determinar la aceptación de uno en vez de otro alegando que esté mejor fundamentado.

— El examen de las operaciones periciales llevadas a cabo por parte de los peritos intervinientes en el proceso, los medios o instrumentos empleados y los datos en los que se sustenten sus dictámenes serán igualmente ponderados, atendiéndose, en el momento de la valoración de los dictámenes, a la competencia profesional de los peritos que suscriban cada informe y a las circunstancias de las que se presuma su objetividad.

De todo lo anteriormente expuesto pueden deducirse dos cuestiones. Por un lado, que no existe como tal un único método científico y que, también en el análisis de patrones de manchas de sangre pueden surgir varias maneras de proceder a la hora de tratar de dar respuesta a un mismo problema. Por otro lado, que, a pesar de la aplicación de esos métodos, que se presuponen objetivos y rigurosos, la disciplina tampoco está exenta de problemas de subjetividad y sesgos que pueden surgir, tanto en la elaboración de los informes como en su interpretación posterior por parte de la administración de justicia.

Así, por poner un ejemplo, la objetividad que lleva aparejada una prueba de orientación para localizar restos hemáticos con un reactivo quimioluminiscente como Luminol, que permite gracias a un proceso químico reproducible en análogas condiciones concretar resultados específicos y demostrables, puede no ser interpretada de manera adecuada si no se cuenta con formación específica, cayendo en errores al valorar un falso positivo. A ello se le ha de sumar la posible subjetividad reinante al momento de analizar los patrones para, por ejemplo, establecer una línea temporal en el episodio sangriento estudiado o concretar el mecanismo que ha podido generar un depósito en cuestión, al no contarse con una guía universal que recoja una única forma de proceder con su clasificación (Hook *et al.*, 2024) y al pesar, en esas circunstancias, la experiencia y formación del analista de forma preponderante. Todo ello se traduce en la posible existencia de sesgos contextuales presentes al momento de describir ciertos elementos por parte del propio perito o durante su valoración e interpretación por parte de otros expertos o de la autoridad judicial.

Resulta interesante en este punto rescatar la relación entre las ciencias forenses y los sesgos que puede afectar a sus investigadores y que han tratado, entre otros, algunos autores como Dror (2020), al afirmar que pueden influir a la opinión de los especialistas en este ámbito, sobre todo, los factores personales y cognitivos, la

capacitación y experiencia y los datos y el material disponibles. Y no solo eso, el propio contexto en el que se desarrolle un caso en particular, y las características de las personas involucradas o la información emocional sobre las circunstancias en las que tiene lugar, pueden evocar atribuciones generalmente basadas en estereotipos en cada especialista (Cippitani *et al.*, 2021). Por no hablar de que, de nuevo ese contexto, configurará la manera en la que se observa el mundo, pudiendo plantearse problemas distintos para cada investigador que partan de una misma cuestión (Toulmin, 1972).

Sesgos que quedarían en gran parte solucionados, al afectar entre otras cuestiones, a la capacitación y experiencia, si se optase por aunar la terminología a emplear y se presentasen estándares sin ambigüedades que diferencien las clasificaciones y descripciones de los patrones en función de su aspecto directo y de las características observables que los conformen, dejando de confiar en posibles inferencias sobre el mecanismo que las genera (Hook *et al.*, 2024) o al momento de concretar la formación y experiencia necesarias para validar la capacidad de cada profesional encargado de desarrollar una pericia o formar parte de una investigación enfocada desde esta disciplina.

En el caso del análisis de patrones de manchas de sangre, la cantidad de información a la que se tiene acceso en cada caso objeto de estudio guarda una relación directa con el nivel de incerteza al que debe enfrentarse el especialista, y, por ende, con las posibilidades de errar en la elaboración y dictamen de sus resultados y conclusiones. Es decir, cuanto más conocimiento tenga el analista, y más herramientas conozca para abordar un mismo problema, menores serán las inferencias de esos sesgos.

Todo ello refuerza la idea de que, en gran parte, el valor de la pericial viene determinada por la formación recibida y la propia pericia del experto, pues permitirán que los razonamientos y conclusiones obtenidas, el método escogido y la información valorada sean adecuados, deduciéndose una cierta predilección por peritos titulares o que puedan demostrar esa formación especializada que se presupone actualizada y adecuada a la cuestión objeto de estudio. Cuestión que genera cierta controversia al denotarse una falta, tanto de criterios para el desarrollo del contenido del documento y, por ende, para la presentación de esos razonamientos, como de formación reglada en el ámbito estudiado que permita demostrar la idoneidad del perfil de cada experto en relación con el objeto de la pericia. Asuntos que quedarán abordados en los siguientes epígrafes y que serán los puntos principales sobre los que asentar las recomendaciones de cara a promover el desarrollo de la disciplina y de las pruebas que deriven de la aplicación de su metodología de la forma más garante y rigurosa posible en pro del propio proceso judicial.

4. ESTADO DEL INFORME PERICIAL EN MANCHAS DE SANGRE EN ESPAÑA

Uno de los puntos más relevantes del estudio recae en la necesidad de comprender el fundamento del informe pericial en manchas de sangre, recogiendo sus características más concretas para poder así cumplir con el objetivo principal: estudiar su

contenido y concretar su protagonismo procesal a fin de aportar reflexiones y posibles mejoras que garanticen su correcto desarrollo. A ello debe sumársele el análisis jurisprudencial, a partir del cual se logrará, no solo conocer casos en los que se haya acudido a esta pericia en la historia más reciente de nuestro país, obteniendo así información sobre su aplicación, sino comprender la forma en la que se procede: quiénes se encargan de realizar el informe, qué estudios se recogen, qué información brindan al proceso, etc., aportando de esta manera valor a esas reflexiones finales, al partir de conocimiento concreto y objetivo sobre el estado de la materia y sus posibles limitaciones o errores.

4.1. Fundamentos de la pericial en manchas de sangre

Como puede deducirse de lo que ha quedado anteriormente recogido, las ciencias forenses han de cumplir con el objetivo principal de investigar, estudiar y analizar los vestigios del hecho criminal concreto para poder formalizar el examen del cuerpo del delito y, con ello, dar lugar a la prueba que instruya el proceso penal (De Barros *et al.*, 2021).

Al contrario de lo que ocurre en la gran mayoría de disciplinas forenses, el estudio de las manchas y patrones desde el punto de vista reconstructivo no será comparativo, al no basarse en el cotejo entre muestra dubitada e indubitada con el fin de establecer una correspondencia. Desde el análisis de patrones de manchas de sangre se procede al estudio de las características de esos depósitos hemáticos para concretar la manera en la que fueron creados, desentrañando, gracias al aspecto que presenten, información de utilidad que permita concretar cómo se fue desarrollando el episodio sangriento en cuestión y, en definitiva, qué ocurrió para cumplir con su objetivo principal: reconstruir el hecho investigado (Hernández Moreno, 2020).

De forma muy general, puede afirmarse que la disciplina y sus análisis se basan en el hecho de que la sangre se adhiere a las leyes físicas de los fluidos, ampliamente predecibles y reproducibles bajo condiciones similares. O lo que es lo mismo, parte de la premisa de que, si se atienden a los resultados de ciertos eventos dinámicos en los que ha habido un derramamiento de sangre, teniendo en cuenta sus propiedades como fluido, se podrán concretar las acciones físicas que dieron lugar a cada mancha o patrón estudiados, lográndose la reconstrucción de cada episodio analizado gracias a la reproducibilidad de esos eventos (Moreno Lopera, 2015).

Para cumplir con su objetivo, y dado que se trata de una disciplina científica, es importante recalcar que los estudios, técnicas y herramientas empleadas desde el análisis de patrones de manchas de sangre con fines reconstructivos han de seguir parámetros y fundamentos concretos, validados por la comunidad internacional de especialistas y que cumplan con el objetivo principal del método científico: permitir que se repitan o reproduzcan de nuevo esos resultados obtenidos. Esto sólo podrá

conseguirse si se acude a esas técnicas concretas y se aboga por procesos científicos detallados y específicos para cada caso, en los que prime el rigor.

De forma concreta, y atendiendo a la información disponible por parte la Asociación Internacional de Patrones de Manchas de Sangre (IABPA), y de otras complementarias como la Asociación Internacional para la Investigación (IAI por sus siglas en inglés), se acepta que deberá nutrirse de otras ciencias como la física, la química o las matemáticas para alcanzar sus objetivos. Así, la mayoría de los análisis desarrollados desde la disciplina precisarán de un conocimiento en profundidad de otras áreas científicas, al centrarse en concretar el área de origen del episodio, empleándose para ello fórmulas trigonométricas; el mecanismo que causó las manchas, debiendo acudir a las propiedades de los fluidos; o en localizar manchas latentes o de difícil visibilidad, empleándose diferentes reactivos químicos en función de cada escenario, mancha o superficie.

Además, y haciendo alusión a ese avance de las técnicas a las que acudir desde la disciplina, deben destacarse la implementación de la tecnología 3D optimizada para la recreación del escenario a través de vídeos o infografías hiper realistas o la mejora en la clasificación de los patrones y manchas de sangre (Amin Tokder Shoukin *et al.*, 2024); las técnicas como la Espectroscopía Raman para concretar la identificación de esas manchas (Almehmadi y Lednev, 2023) o para determinar su antigüedad (Barber *et al.*, 2024); o la utilización de ecuaciones físicas para predecir la altura y el volumen de las gotas de sangre encontradas en un lugar concreto (Ukanchanakitti *et al.*, 2023).

Avances que ponen de manifiesto la necesidad de que los expertos encargados de proceder con la realización y ratificación de los informes periciales en esta disciplina, además de contar con estudios específicos sobre sus fundamentos, mantengan una formación continua, atendiendo a los cambios y mejoras que, año tras año, quedan aceptadas por la Asociación Internacional e implementadas en la comunidad de expertos.

4.2. Contenido del informe pericial

El informe pericial, tal y como recoge la LECrim en el artículo 475, deberá ceñirse al objeto de estudio que se hubiera facilitado por el juez de instrucción, debiendo elaborarse su contenido y las conclusiones correspondientes siguiendo siempre ese tema en concreto (Montero Aroca, 2019). De forma más específica, en el artículo 478 de esa misma ley, se incide en la importancia de que, siempre que fuera posible, el documento pericial deba comprender una descripción de la persona o cosa que sea objeto del informe en el estado o modo en el que se halle; una relación detallada de todas las operaciones que se practiquen por los peritos así como de su resultado, y las conclusiones que formulen los peritos en función de los datos que obtengan siguiendo los principios de su ciencia o arte. Más allá de estos apuntes, lo cierto que es que

hay cierta escasez de legislación y normativa jurídica al respecto en nuestro país, muy probablemente debido a la enorme cantidad de ciencias, artes y técnicas sobre las que puede versar el informe, lo que impide la posibilidad de abordar una legislación estricta que recoja qué debe o qué no debe incluir cada dictamen. Se presume, por tanto, que el informe pericial deberá regirse en función de cada especialidad, atendiendo a las recomendaciones de la LECrim y a las propias técnicas y métodos de estas.

Ahora bien, como se ha venido comentando, esta falta de información, criterios y normativa concretas en nuestro país no debe aceptarse como la tónica general dentro de la disciplina, pues, en otros países, sí se cuenta, además de con formación específica y reglada, con guías y estándares para la correcta redacción del informe pericial en patrones de manchas de sangre. Un ejemplo de ello es el modelo que se recoge desde el *National Institute of Standards and Technology* (NIST), preparado, aceptado y aprobado por la *Academy Standards Board* (ASB) en colaboración con la *Organization of Scientific Area Committees for Forensic Science* (OSAC, 2017) y que se basa en un documento originado anteriormente, en 2009, por el grupo SWGSTAIN (*Scientific Working Group on Bloodstain Pattern Analysis*), altamente cualificado en esta materia (SWGSTAIN, 2009).

En él se presentan, a modo de guía, los apartados que debería incluir un informe pericial en manchas de sangre, independientemente de que sean estudios sobre escenarios, análisis de imágenes o vídeos o exámenes de evidencias. La gran mayoría de ellos quedan recogidos en nuestra LECrim (art. 478)⁵, aunque, en esta ocasión, las indicaciones son mucho más detalladas y concretas, llegándose a presentar ejemplos de información que podría añadirse en cada epígrafe o del lenguaje recomendado para presentar ciertas afirmaciones o conclusiones. Además, se hace referencia a la importancia de incluir el área de especialización del autor para que pueda valorarse, de forma concreta, el contenido.

De esta manera, si se establecen apreciaciones que exceden sus competencias, estas deberán ser desestimadas al no formar parte del área de estudio del perito y, por tanto, no contar estar respaldadas de forma rigurosa o veraz. Mismo objetivo persigue la presentación de las técnicas de forma adecuada, la inclusión de información sobre la localización concreta de las muestras estudiadas, los documentos analizados o los materiales revisados durante el estudio, o la explicación de posibles limitaciones durante el proceso en relación con alguna técnica empleada o referidas a las propias

⁵ El artículo hace referencia a algunos elementos tales como la descripción de la persona o cosa objeto del mismo en el estado o modo en que se halle; la relación detallada de todas las operaciones practicadas por los peritos y de su resultado, extendida y autorizada en la misma forma que la anterior y de las conclusiones que en vista de tales formulen los peritos conforme a los principios y reglas de su ciencia o arte. Además de ello, el estándar ANSI/ASB (2020) al que se alude recoge, en su primera edición, la necesidad de atender a otros elementos en los informes de la disciplina tales como: los métodos empleados en el examen del elemento; la localización donde se produjeron esos exámenes (cuando no los llevasen a cabo los especialistas en sus laboratorios); las limitaciones; el procedimiento de toma de muestras o la declaración de cualificación del perito, entre otras.

muestras (OSAC, 2017). Limitaciones que pueden ser relativamente comunes si tenemos en cuenta la fragilidad del vestigio objeto de estudio, de carácter biológico y especialmente fácil de degradar y contaminar, sobre todo, si se encuentra en escenarios abiertos, sin protección estructural y a merced de fenómenos meteorológicos, por ejemplo.

Con este tipo de guías, lo que se pretende es evitar el menoscabo de la disciplina, velando por su correcto desarrollo al tratar de evitar que, profesionales de otros campos sin la formación adecuada, puedan inmiscuirse en un área de especialización que no es la suya, con las limitaciones de conocimiento que conlleva y el desprestigio que puede suponer, más aún si tenemos en cuenta que se trata de una disciplina científica de la que, como tal, se espera un rigor absoluto. Formación que, si bien no excluye la existencia de los posibles sesgos contextuales que puedan someter a la disciplina a un estado de vulnerabilidad, tal y como se comentó anteriormente, contribuye a minimizarlos. Al exigirse la adopción de protocolos y métodos concretos, se reduce la posibilidad de optar por una técnica, herramienta o forma de proceder en desuso, poco concreta o subjetiva, reduciendo posibles errores, tanto en el momento de analizar las muestras y redactar el informe, como durante la valoración del mismo por parte de la autoridad judicial.

4.3. Revisión jurisprudencial

Como quedó anteriormente recogido es importante, para poder tener una visión generalizada del estado de la disciplina y, sobre todo, de su aplicación real, proceder con un análisis jurisprudencial que permita concretar la forma en la que se introduce el informe pericial de manchas de sangre en el proceso, cuando así ocurre, permitiendo con ello atender a sus limitaciones y presentar y valorar posibles puntos de mejora.

En líneas generales, y como quedará más adelante expuesto, apenas se han presentado informes periciales sobre la materia en nuestro país pues, de forma generalizada, el estudio de la sangre ha quedado circunscrita a fines identificativos, presentándose estudios de ADN. De hecho, en la gran mayoría de las ocasiones en las que se produce un delito violento en el que se atenta contra la vida o la integridad física, y la sangre adquiere cierto protagonismo, no se suele recurrir a la disciplina para lograr más información, apostando por el estudio de otros vestigios como las huellas de pisada, los residuos de disparo o las huellas dactilares para recopilar información relevante sobre lo ocurrido.

En los casos encontrados, todos ellos pertenecientes a un periodo temporal que comprende desde el año 2006 al 2019, llama la atención, tal y como quedará desgranado a continuación, la falta de información concreta sobre el propio informe, su contenido y el perfil del perito. Es más, de forma mayoritaria, no llega a presentarse ni siquiera el documento de manera formal; se acepta cierta información basándose en la deducción lógica o en la presencia o ausencia de sangre en ciertos lugares,

resultando realmente complejo proceder con un análisis real del informe pericial en manchas de sangre al no contar con datos sobre su contenido, la forma en la que se presenta o la información que aporta al propio proceso, permitiendo establecer ya una primera limitación: la falta efectiva de periciales sobre la disciplina.

Así, en varias de las sentencias consultadas⁶, la información sobre las manchas de sangre que se recoge en el acta de la Inspección Técnico Ocular (ITO), desarrollada por miembros de la policía científica, fue suficiente como para auxiliar al tribunal en la toma de decisiones durante el proceso judicial, al valorarse la posibilidad de que las máculas observadas en la escena guardaran relación con los hechos descritos de forma lógica y, por tanto, fueran obvias en la comisión de los mismos, recurriéndose al análisis de ADN como prueba de cargo para la resolución de cada caso, pero no al análisis de patrones de manchas de sangre para su reconstrucción. Simplemente se aportan

⁶ STS 135/2017, de 2 marzo 2017. No se estudia el análisis de la sangre encontrada en la escena para proceder a la resolución del caso. Se establece la existencia de huellas de pisada ensangrentadas, de un charco de sangre en la escena y junto al cuerpo de la víctima y en su ropa.

A pesar de ello, se llevan a cabo únicamente estudios para determinar la presencia de sangre en unos pantalones con la intención de llevar a cabo pruebas de ADN.

SAP de Cádiz, 310/2016 de 27 de octubre de 2016. Se recoge la presencia, en un caso de doble homicidio con arma blanca, de sangre de las dos víctimas, tanto en el dormitorio en el que se encontraba una de ellas, como en un pasillo en el que se encontraba la otra víctima. Se realizan estudios de ADN, pero no periciales sobre esas manchas de sangre.

SAP de Huesca 89/2016 de 12 de julio de 2016. Se describe la presencia de sangre en varias zonas de la escena, pero no se lleva a cabo ninguna pericia sobre las manchas en cuestión.

STS 844/2017 de 21 de diciembre de 2017. Caso de intento de homicidio. Se recoge la presencia de un gran charco de sangre, salpicaduras y manchas por transferencia en varias prendas. Sin embargo, se realizaron exclusivamente análisis de ADN sobre esa sangre. Las decisiones del tribunal se apoyaron en varios aspectos presentados en el acta de la IOTP, entendiéndose que no casaban las declaraciones del acusado con lo observado en la escena.

STS 136/2012 de 6 de marzo de 2012. Se estudia la presencia de la sangre en la escena, y en la ropa del acusado, para auxiliar la toma de decisiones del tribunal, al entender que tanto las huellas ensangrentadas encontradas, como las pequeñas gotas recogidas en la camisa del acusado, no guardaban relación con las declaraciones que había brindado a los investigadores. A pesar de ello, no se recoge la realización de ninguna pericia. Únicamente se establece una relación entre la sangre hallada en la IOTP, la declaración del acusado, el examen patológico a la víctima y los resultados del análisis de ADN de parte de esa sangre.

STS 366/2005 de 28 de marzo de 2005. Se deduce la presencia de manchas de sangre en una prenda con la imposibilidad de que la declaración del acusado sea veraz, pero tampoco se llevó a cabo una pericial especializada en la materia.

SAP de Madrid 134/2016 de 17 de marzo de 2016. Se establece la presencia de diferentes manchas de sangre y huellas, tanto de pisada como dactilares también con sangre, en el acta de la IOTP. A lo largo de la sentencia, al referirse a la presencia de manchas y resto de sangre se establece en varias ocasiones «se cree que», «podría deberse a», «puede ser compatible». Debería haberse acudido a un estudio de manchas de sangre y de localización de estas en la escena para evitar que pudieran darse fallos en su estudio al no contarse con los conocimientos suficientes como para discernir unas manchas de otras a simple vista y sin un examen en profundidad.

Al igual que en las sentencias anteriores, sí se acude a informes de huellas de pisada o dactilares, además del análisis de ADN.

resultados sobre la relación lógica entre la presencia de la sangre en ciertas zonas de la escena y los datos que aportan la ITO y el resto de los vestigios, acompañándolo con resultados de pruebas de ADN, para enfatizar las conclusiones en cada sentencia.

De forma más específica, en uno de los casos estudiados (SAP Las Palmas de Gran Canaria 27/2013 de 23 de abril de 2013 (recurso 2/2012) se recurre a la elaboración de un informe por parte de un médico forense en colaboración con miembros de criminalística de la Policía Nacional con el objetivo de determinar si se había producido un homicidio atendiendo de forma exclusiva a la sangre encontrada en la escena criminal. En comparación con otras sentencias, se trata del caso que mejor y de forma más específica ha sabido incluir el análisis de manchas de sangre en el procedimiento hasta este momento, pero sigue presentando ciertas limitaciones. Y es que, a pesar de que se recoge que la médico forense realiza el análisis basándose en bibliografía científica y con el apoyo de los miembros de la policía científica, se desconoce si cuenta con los conocimientos especializados que le permiten ponerse al frente de este tipo de pericial y tener acceso a los métodos y técnicas específicas para la manipulación y el tratamiento de las muestras de sangre para su posterior análisis de forma garante.

En otra de las sentencias consultadas se recoge que:

«en base a la inspección técnico ocular de la Guardia Civil, obrante a los ff 92 y ss., al informe sobre ropa y efectos de la víctima; por la pericial de Criminalística obrante a los ff 259 y 316, que con fecha 17.06.15 relata que la ausencia de marcas de manchas de proyección dinámica y con la pericial de 27.02.17 obrante al f 523 en que se indica que la sangre del acusado lo es porque el acusado tendría que estar al lado de la víctima, porque rápidamente se da cuenta de lo que ha hecho y coge la cabeza y se la pone en el pantalón, se determina que todo indica que cuando se produce la herida mortal el acusado tendría que haber estado muy cerca» (STS 1/2019, de 10 de enero de 2019, p. 6).

A diferencia de los casos anteriores, en esta sentencia sí se menta la realización de un informe pericial específico para determinar la lógica de la presencia o ausencia de las manchas de sangre en la escena, realizándose un estudio en el que se tienen en cuenta, tanto el acta de la ITO en el que se recogió la presencia de sangre en la escena, como un análisis posterior sobre las prendas que la víctima vestía en ese momento para reconstruir el episodio sangriento y determinar las posiciones y movimientos de aquella y del agresor en cada momento, aportando objetividad y datos veraces fácilmente comprobables al proceso judicial. Se presupone, además, dada la especificidad del contenido, que los especialistas encargados de su realización contaban con formación concreta y relativa al objeto de estudio, a pesar de que no se logra acceder a esa información.

Mención aparte merece el papel de esta disciplina en el caso del doble asesinato en Almonte, Huelva, en el año 2013⁷. Lo complejo de la investigación, unido a la

⁷ El 27 de abril de 2013 un padre y su hija de ocho años recibían más de 150 puñaladas en su domicilio, descubriéndose los cadáveres por un familiar un par de días después.

falta de nueva información y/o de vestigios que incriminaran a alguno de los sospechosos de forma determinante, llevó a las partes a solicitar la participación de un perito especialista para que realizase un análisis de las manchas de sangre que se habían encontrado en la escena y pudiera, con ello, arrojar algo de luz al proceso que había quedado estancado. La inclusión del informe pericial en el que se estudiaron los patrones localizados en varias estancias del apartamento aportó datos objetivos, y hasta el momento desconocidos. Se establecieron, entre otros aspectos, cinco posibles áreas de origen de la fuente de sangrado (tras realizar los cálculos trigonométricos apropiados), lo que permitió determinar las posiciones de cada víctima y, por tanto, del agresor en el momento en el que los atacó, aportando información de los posibles movimientos de los tres en cada estancia y en cada momento del episodio sangriento, al dibujarse una línea temporal que armaba una hipótesis objetiva y veraz sobre lo ocurrido (Moreno Lopera, 2018; Europa Press Andalucía, 2017).

Sin embargo, y a pesar de lo valioso de este último caso, y considerando lo anteriormente expuesto en relación a la lectura y comprensión de las sentencias, se evidencia el desconocimiento por parte de los investigadores y de los miembros de los tribunales, tanto de la disciplina del análisis de patrones de manchas de sangre, como de la eficacia de los informes que pueden desarrollarse en ese ámbito, revelando además un verdadero problema, pues en muchos de los casos consultados, la realización de esta pericia podría haber arrojado información nueva y, sobre todo, valiosa sobre los hechos que se juzgaban, y jamás se solicitó esta diligencia.

5. REFLEXIONES Y RECOMENDACIONES DE MEJORA.

En sentido estricto, el informe pericial no es sino un documento que recoge un estudio realizado por parte de un experto en un área técnica, científica, artística o práctica sobre la que hay controversia, incluyéndose entre los medios de prueba que pueden presentarse en un proceso (Montero Aroca, 2019) y quedando amparado y protegido jurídicamente por diferentes normas, siendo numerosos los preceptos que regulan su desarrollo (Abel Lluch y Picó i Junoy, 2009; Cameriere *et al.*, 2013; Montero Aroca, 2019).

Entre esa normativa destaca la LECrim, que recogerá su regulación en la fase de instrucción en los artículos 456 al 458, en la fase de instrucción en los artículos 661 al 663 y en el juicio oral en los artículos 723 al 725. Quedando, además, ampliamente recogido como diligencia de investigación y como medio necesario del auxilio judicial en los artículos 465 al 485 de esa misma ley, poniendo así de manifiesto la importancia de la colaboración de la figura del perito en el proceso; más aún a raíz del avance que han experimentado durante las últimas décadas las ciencias forenses y sus diversas disciplinas como el análisis de patrones de manchas de sangre que, si bien resulta relevante gracias a la información que puede aportar para la reconstrucción de los hechos, desde un prisma científico, veraz y riguroso, no cuenta, todavía, ni con el reconocimiento ni con el respaldo normativo que debiera.

Los puntos tratados en el presente artículo ponen de manifiesto la necesidad de contar con pautas de actuación que permitan, no solo regular la formación para poder contar con peritos expertos en la materia que puedan proceder con el desempeño de su labor de forma garante, sino establecer protocolos que les guíen en el desarrollo de sus labores profesionales con la intención de evitar sesgos y subjetividad al momento de desempeñar sus funciones.

En base a ello sería conveniente, por un lado, apostar por la adopción de la terminología sugerida desde la ASB para la identificación y descripción de cada formación sanguínea⁸, así como de las clasificaciones taxonómicas propuestas por la IABPA, más específicas y concretas, y con las que se trata de impedir acudir a otras fundamentadas en los mecanismos que crean los patrones, al poder generar conclusiones subjetivas y sesgos contextuales.

Se aconseja, además, que estos protocolos específicos sean extensivos y hagan alusión, siguiendo igualmente las pautas de la ASB, a las partes que deben quedar cumplimentadas y desarrolladas en un informe pericial⁹, homogeneizando sus actividades y el contenido de los dictámenes sobre análisis de patrones de manchas de sangre que pudieran formar parte de los procesos judiciales en España.

Igualmente interesante resultan las recomendaciones recogidas desde la Academia Nacional de Ciencias Forenses (NAS) (Cippitani *et al.*, 2021) de cara a comprobar la fiabilidad y los límites de rendimiento en diversas disciplinas forenses, centradas en valorar y evaluar su metodología para evitar en la medida de lo posible una alta tasa de error al basarse en «conjeturas subjetivas» o presentar «términos generalmente categóricos» (McQuiston-Surrett y Saks, 2008, p. 1189). Premisa que se ha comentado anteriormente al inferir, por ejemplo, en la necesidad de actualizar metodologías siguiendo las tendencias actuales o de tratar de minimizar la subjetividad en ciertos análisis evitando recurrir a clasificaciones obsoletas y sesgadas.

Tras el análisis jurisprudencial resulta evidente que la falta de informes periciales que versen sobre esta disciplina en nuestro país, así como las limitaciones reinantes en ellos en los casos encontrados al no quedar completamente claro ni el tipo de pericia realizada de forma concreta ni el perfil de quién lo firmó, pueden poner en riesgo la veracidad de la materia y de los resultados aportados al proceso.

La aplicación y adopción de estos protocolos permitiría, al centrar las pautas por las que deben regirse todos los especialistas, tanto durante las formaciones, como

⁸ El informe técnico 033 de la ASB (2017) recoge un listado con los 37 términos recomendados al momento de designar una formación sanguínea que deberían emplearse en publicaciones, manuscritos, en las conclusiones y las discusiones de informes forenses, durante una ratificación o declaración en juicio y en cualquier curso en el que se enseñe sobre análisis de patrones de manchas de sangre.

⁹ La primera edición del estándar 031 (ANSI/ASB, 2020) recoge una serie de guías y pautas para la redacción de un informe pericial en la materia, indicándose igualmente ciertas afirmaciones que deben evitarse. Además de las partes que se aconseja incluir en el documento, se establecen breves instrucciones sobre qué debe redactarse en cada uno de ellos.

durante el desarrollo de su actividad profesional posterior, una mejora en la difusión y presentación de los resultados de sus investigaciones, de la redacción y ratificación de los informes periciales y del resto de asuntos relacionados con la disciplina (divulgación, colaboración con otras disciplinas, etc.), facilitando con ello, igualmente, su comprensión por parte de la autoridad judicial competente.

Resulta indispensable, por último, lograr el reconocimiento de la formación y experiencia de sus profesionales, debiendo expedirse, como ocurre en otros países europeos como Países Bajos, el nivel de *experto* por parte de la autoridad judicial, evitando con ello el intrusismo laboral que pudiera desprestigiar la disciplina, y aunar los itinerarios formativos de los cursos de especialización que se imparten a las diferentes unidades policiales de nuestro país, unificando procesos, metodología y terminología para todos los analistas que ejerzan en su labor en España.

Por todo lo anterior, debe afirmarse que es preciso contar con especialistas formados en la materia que colaboren en el desarrollo de protocolos y métodos de actuación específicos que minimicen las limitaciones referidas a una mala praxis durante el desarrollo de su labor, tanto en la escena como en el laboratorio, garantizando así una inclusión de la pericial en manchas de sangre de manera objetiva y rigurosa en el proceso judicial.

BIBLIOGRAFÍA

- Abel Lluçh, X. y Picó i Junoy, J. (2009). *La prueba pericial*. Bosch Editor.
- Almehmadi, L., y Lednev, I. (2023). Stand-off Raman spectroscopy is a promising approach for the detection and identification of bloodstains for forensic purposes. *Journal of Raman Spectroscopy*, 55(2), 227-231. <https://doi.org/10.1002/jrs.6609>
- Amin Todker Shoukin, M.A., Tasmia, J., Hossain, M.A., y Haque, N. (2024). Bloodstain Classification in Forensic Analysis Using Optimized 3D CNN. En Institute of Electrical and Electronics Engineers [IEEE] (ed.), *6th International Conference on Electrical Engineering and Information and Communication Technology (ICEEICT)* (p. 1246-1251). <https://doi.org/10.1109/ICEEICT62016.2024.10534494>
- Andersen, H. y Hepburn, B. (2015). Scientific Method. En *Stanford Encyclopedia of Philosophy*. <https://plato.stanford.edu/archives/win2020/entries/scientific-method/>
- ANSI/ASB (2020). *ANSI/ASB Standard 031: Standard for Report Writing in Bloodstain Pattern Analysis* (First Edition). https://www.aafs.org/sites/default/files/medial/documents/031_Std_e1.pdf
- ASB (2017). *ASB Technical Report 033: Terms and Definitions in Bloodstains Pattern Analysis* (First Edition). https://www.aafs.org/sites/default/files/medial/documents/033_TR_e1_2017.pdf
- Asencio Cabot, E. (2014). Una aproximación a la concepción de ciencia en la contemporaneidad desde la perspectiva de la educación científica. *Ciencia y Educação*, 20(3), 549-560 <https://doi.org/10.1590/1516-73132014000300003>
- Barber, A. P., Weber, A. R. y Lednev, I. K. (2024). Raman spectroscopy to determine the time since deposition of heated bloodstains. *Forensic Chemistry*, 37. <https://doi.org/10.1016/j.forc.2024.100549>
- Bevel, T. y Gardner, R. M. (2008). *Bloodstain Pattern Analysis with an Introduction to Crime Scene Reconstruction* (3rd edition). CRC Press.
- Cameriere, R., De Luca, S., y Navarro, F. (2013). La prueba pericial y su valoración en el ámbito judicial español. *Revista electrónica de ciencia penal y criminología* (15-19). <http://criminnet.ugr.es/recpc/15/recpc15-19.pdf>

- Cando Shevchukova, S. L. (2023). La prueba grafotécnica: fundamentos, validez y fiabilidad. *Quaestio Facti. Revista Internacional Sobre Razonamiento Probatorio*, 4, 275–303. https://doi.org/10.33115/udg_bib/qf.i.22841
- Cippitani, R., Mirabile A., y Onofri, M. (2021). «Objetividad científica» y sesgos en la toma de decisiones jurídicas: casos de genética forense y de algoritmos. *Revista Justicia y Derecho*, 4(2), 1-13. <https://doi.org/10.32457/rjyd.v4i2.1440>
- De Barros, F., Kuhnen, B., Da Costa Serra, M. y Da Silva Fernandes, C. (2021). Ciencias Forenses: principios éticos y sesgos. *Revista Bioética*, 29(1), 55-65. <https://doi.org/10.1590/1983-80422021291446>
- Decisión marco 2009/905/JAI del Consejo, de 30 de noviembre de 2009, sobre acreditación de prestadores de servicios forenses que llevan a cabo actividades de laboratorio. *DOUE*, núm 322, de 9 de diciembre de 2009. <https://www.boe.es/buscar/doc.php?id=DOUE-L-2009-82381>
- Dror, I. (2020). Cognitive and Human Factors in Expert Decision Making: Six Fallacies and the Eight Sources of Bias. *Analytical Chemistry*, 92(12). <https://dx.doi.org/10.1021/acs.analchem.0c00704>.
- Egg, A. (1977). *Introducción a las técnicas de investigación social*. Humanitas.
- Europa Press Andalucía (19 de septiembre de 2017). Peritos destacan el episodio brutal en la casa del crimen de Almonte y creen que conocía la vivienda. *Europa Press* <https://www.europapress.es/andalucia/huelva-00354/noticia-peritos-destacan-episodio-brutal-casa-crimen-almonte-huelva-creen-autor-conocia-vivienda-20170919144149.html>
- Hook E., Fieldhouse S., Flatman-Fairs D., y Williams G. (2024). Bloodstain classification methods: A critical review and a look to the future. *Science & Justice*, 64(4), 408-420. <https://doi.org/10.1016/j.scijus.2024.06.004>
- Feynman, R. (2001). ¿Qué es la ciencia? *Polis: Revista Latinoamericana*, 1. <https://journals.openedition.org/polis/8245>
- Gascón Abellán, M. (2020). Prevención y educación: el camino hacia una mejor ciencia forense en el sistema de justicia. En J. Ferrer Beltrán y C. Vázquez (Eds.), *El razonamiento probatorio en el proceso judicial. Un encuentro entre diferentes tradiciones*. Marcial Pons.
- Gonáni, O. A. (2012). La prueba científica no es prueba pericial. *Derecho y Sociedad*, 38, 169-175. <https://revistas.pucp.edu.pe/index.php/derechosociedad/article/view/13116>.
- Hernández Moreno, M. (2020). *Test orientativos para la localización de manchas de sangre. Estudio comparativo y aplicación en la escena del delito*. [Tesis Doctoral, Universidad de Alcalá].
- Hernández Moreno, M. (2022). Bloodstain pattern analysis as scientific evidence in the judicial process. State of the matter in Spain and lege ferenda proposal. *Spanish Journal of Legislative Studies*, 4, 1-17. <https://doi.org/10.21134/sjls.vi4.1743>
- James, S. H. y Eckert, W. G. (1999). *Interpretation of bloodstain evidence at crime scenes* (2nd edition). CRC Press
- James, S. H., Kish, P. E. y Sutton, T. P. (2005). *Principles of bloodstain pattern analysis: Theory and practice*. CRC Press Edition.
- Illes, M. (2001). Canadian bloodstain pattern analysis in the Netherlands. *Canadian Society of Forensic Science Journal*, 34(4), 167-171. <https://doi.org/10.1080/00085030.2001.10757526>
- McQuiston-Surrett, D. y Saks, M. J. (2008). Communicating Opinion Evidence in the Forensic Identification Sciences: Accuracy and Impact. *Hastings Law Journal*, 59(5), 1159-1189. https://repository.uclawf.edu/hastings_law_journal/vol59/iss5/71
- Meneses Pacheco, C. (2008). Fuentes de prueba y medios de prueba en el proceso civil. *Ius Et Praxis*, 14(2), 43-86. <http://dx.doi.org/10.4067/S0718-00122008000200003>
- Meza, M. (2024). El informe pericial químico en la determinación de residuales de disparo. Contenidos mínimos, razonamiento experto y debido proceso. *Quaestio Facti. Revista Internacional Sobre Razonamiento Probatorio*, 6, 285–298. https://doi.org/10.33115/udg_bib/qf.i6.22982
- Montero Aroca, J. (2019). *Derecho jurisdiccional* (Volumen III: Proceso penal, 27ª ed.). Tirant lo Blanch.
- Moreno Lopera, M. (2015). *Análisis e interpretación de los patrones de manchas de sangre: Experiencia, conocimientos, formación y opiniones de los profesionales implicados en la investigación criminalística*. [Tesis Doctoral, Universidad de Murcia]. Digitum: repositorio institucional de la Universidad de Murcia. <http://hdl.handle.net/10201149703>

- Moreno Lopera, M. (2018). Análisis e interpretación de los patrones de manchas de sangre. Estudio y reconstrucción. *Minerva: Saber, Arte y Técnica*, 2(2), 6-23. <https://ojs.editorialiupfa.com/index.php/minerva/article/view/87>
- Organization of Scientific Area Committees for Forensic Science [OSAC] (2017). *Guidelines for Report Writing in Bloodstain Pattern Analysis*. https://www.nist.gov/system/files/documents/2017/11/15/bpa_guidelines_for_report_writing_in_bloodstain_pattern_analysis.pdf
- Peyrano J. W. (2015). Sobre la prueba científica. *Ius: la revista*, 35, 108-113. <https://revistas.pucp.edu.pe/index.php/iusetveritas/article/download/12287/12851/0>
- Sánchez Rubio, A. (2016). *Ciencia y proceso penal. un estudio sobre el concepto y régimen jurídico de la llamada «prueba científica»*. [Tesis Doctoral, Universidad Pablo de Olavide].
- Scientific Working Group on Bloodstain Pattern Analysis [SWGSTAIN] (2009). Recommended Terminology. *Forensic Science Communications*, 11(2), 1-5. https://theiai.org/docs/SWGSTAIN_Terminology.pdf
- Toulmin, T. (1972). *Human Understanding*. Clarendon Press.
- Ukanchanakitti, P., Winaichatsak, N., Cho, N. y Sumetpipat., K. (2023). Predicting Blood Drop Height and Volume using Physics Equations, VGG-19, and XGBoost. *LAIT '23: Proceedings of the 13th International Conference on Advances in Information Technology*, 22, 1-8. <https://doi.org/10.1145/3628454.363115>
- Vargas Ávila, R. (2010). La valoración de la prueba científica de ADN en el proceso penal. *Prolegómenos: derechos y valores*, 13(25), 127-146. <https://doi.org/10.18359/prole.2450>
- Vázquez, C. (2014). Sobre la científicidad de la prueba científica en el proceso judicial. *Anuario de Psicología Jurídica*, 24(1), 65-73. <https://doi.org/10.1016/j.apj.2014.09.001>
- Vázquez, C. (2015). *De la prueba científica a la prueba pericial*. Marcial Pons.

Sentencias

- España. Real Decreto de 14 de septiembre de 1882 por el que se aprueba la Ley de Enjuiciamiento Criminal. (Gaceta de Madrid, núm. 260, de 17/09/1882). <https://www.boe.es/eli/es/rd/1882/09/14/1/con>
- Sentencia de la Audiencia Provincial de Cádiz (España) 310/2016 (Sala de lo Penal), de 27 de octubre de 2016. ES:APCA:2016:1528
- Sentencia de la Audiencia Provincial de Huesca (España) 89/2016, de 12 de julio de 2016. ES:APHU:2016:87
- Sentencia de la Audiencia Provincial de Las Palmas de Gran Canaria (España) 27/2013, de 23 de abril de 2013. <https://vlex.es/vid/435065714>
- Sentencia de la Audiencia Provincial de Madrid (España) 134/2016, de 17 de marzo de 2016. <https://www.legaltoday.com/historico/jurisprudencial/jurisprudencia-penal/sentencia-num-1342016-audiencia-provincial-madrid-17-03-2016-2016-05-09/>
- Sentencia del Tribunal Supremo (España) 366/2005 (Sala de lo Penal), de 28 de marzo de 2005. ECLI:ES:TS:2005:1868
- Sentencia del Tribunal Supremo (España) 136/2012 (Sala de lo Penal), de 6 de marzo de 2012. ECLI:ES:TS:2012:1699
- Sentencia del Tribunal Supremo (España) 844/2017 (Sala de lo Penal), de 21 de diciembre de 2017. ECLI:ES:TS:2017:4667
- Sentencia del Tribunal Supremo (España) 1/2019 (Sala de lo Penal), de 10 de enero de 2019 (recurso 10388/2018). ECLI:ES:TS:2019:237.
- Sentencia del Tribunal Supremo (España) 135/2017, de 2 de marzo de 2017. <https://vlex.es/vid/671290689>

LA CAUSALIDAD Y LA *NORMATIVIDAD* EN LA PRUEBA TRIBUTARIA. REFLEXIONES A PARTIR DE LA REGLA DE EXCLUSIÓN PROBATORIA

Serrana Delgado Manteiga*

RESUMEN: El presente artículo aborda el problema de la regla de exclusión probatoria cuando se aplica al caso de la prueba tributaria. Enfatiza en que, dado que las Administraciones Tributarias desarrollan procedimientos de obtención de prueba para los que gozan de prerrogativas especiales que no tienen los particulares, se impone distinguir entre la causalidad y la normatividad de la prueba. Basándome en los conceptos de reglas constitutivas e institucionales y en la distinción entre acción y descripción de la acción, concluyo en que un mismo hecho probatorio puede ser una prueba lícita o ilícita según el marco institucional en el que se inserta y que es necesario distinguir conexión causal de conexión normativa (entendida esta última como algo diferente a la conexión de antijuridicidad que construyó la dogmática y la jurisprudencia).

PALABRAS CLAVE: regla de exclusión; causalidad; normatividad; reglas constitutivas; descripción de acción.

CAUSALITY AND NORMATIVITY IN TAX EVIDENCE. REFLECTIONS BASED ON EXCLUSIONARY RULE

* Profesora Adjunta de Filosofía y Teoría General del Derecho de la Facultad de Derecho de la UDELAR y es Profesora de Marco Constitucional y Marco Legal de la Maestría en Tributación de la Facultad de Ciencias Económicas de la UDELAR. Las opiniones vertidas en el presente artículo son de exclusiva responsabilidad de la autora y no comprometen a ninguna institución.

ABSTRACT: This article addresses the problem of the exclusionary rule when applied to the case of tax evidence. It emphasizes that, given that Tax Administrations develop procedures for obtaining evidence for those who enjoy special prerogatives that individuals do not have, it is necessary to distinguish between causation and the *normativity* of evidence. Based on the concepts of constitutive and institutional rules and on the distinction between action and description of the action, I conclude that the same evidentiary fact can be licit or illicit evidence depending on the institutional framework in which it is inserted and that it is necessary to distinguish causal connection from normative connection (the latter understood as something different from the connection of illegality that dogmatics and jurisprudence constructed).

KEYWORDS: exclusionary rule; causation; normativity; constitutive rules; description of the action.

SUMARIO: 1. INTRODUCCIÓN.— 2. LA TENSIÓN ENTRE LA VERDAD Y LAS GARANTÍAS EN EL CASO DE LA PRUEBA TRIBUTARIA.— 3. LOS CASOS TRIBUTARIOS DILEMÁTICOS Y LAS REGLAS DE EXCLUSIÓN PROBATORIA: 3.1 El caso Falciani; 3.2 El caso del frigorífico uruguayo.— 4. EL PROBLEMÁTICO CONCEPTO DE HECHO EN EL DERECHO: 4.1 Los hechos y las reglas constitutivas o institucionales; 4.2 Los hechos y los niveles de lenguaje.— 5. LA AMBIGÜEDAD EN EL TÉRMINO PRUEBA Y PRUEBA ÍLÍCITA.— 6. LA DIFERENCIA ENTRE LA CAUSALIDAD Y LA NORMATIVIDAD Y LA REGLA DE EXCLUSIÓN PROBATORIA: 6.1 Las acciones y las descripciones de las acciones: el intento naturalista de Davidson que coadyuva a reafirmar la distinción entre causalidad y normatividad; 6.2 La causalidad y la normatividad en las reglas de exclusión probatoria a la luz de los casos tributarios bajo estudio.— 7. CONCLUSIONES.— BIBLIOGRAFÍA.

1. INTRODUCCIÓN

El objetivo del presente trabajo es problematizar en torno a la regla de exclusión probatoria, específicamente la tesis del fruto del árbol envenenado, distinguiendo las categorías conceptuales de *causalidad* y *normatividad* en el ámbito de la teoría de la prueba. Me propongo demostrar que la dogmática, la jurisprudencia y ciertas concepciones de la teoría de la prueba no han distinguido estos diferentes ámbitos de discurso racional, lo que ha llevado a una pérdida de claridad conceptual para realizar análisis justificados de los problemas involucrados. Tomo como ejemplo dos sentencias que refieren a la materia tributaria (sección 3), una española y otra uruguaya, donde ambos tribunales llegan a conclusiones contradictorias en casos muy similares, pero ambos comparten la no distinción entre los diferentes niveles de análisis: el causal por un lado y el normativo por el otro.

En la sección 2 sostengo que la prueba tributaria no presenta particularidades que hagan necesario un análisis conceptual diferente al de otras materias, pero que puede haber algunas cuestiones de énfasis que valga la pena resaltar, entre ellas, que las Administraciones Tributarias tienen procedimientos administrativos para la búsqueda de la verdad a la que están orientadas, lo que resulta relevante para la tesis sostenida en el artículo.

En la sección 4 distingo las categorías conceptuales de hecho bruto y hecho institucional. Me valgo de la definición tarskiana de verdad y de la noción de regla constitutiva de Searle (1995). Introduzco la definición de verdad de Tarski (1969) porque esta es útil para distinguir niveles de lenguaje, donde la noción de verdad no es ajena a la idea de que puede involucrar la correlación con una definición dada o acordada en un nivel inferior de lenguaje. Se trata de una definición recursiva. Esto implica que la verdad es la correspondencia de un enunciado con otro enunciado situado en un nivel de lenguaje inferior. Esta tesis es compatible con una visión no escéptica en materia de verdad y de prueba y contribuye a esclarecer como un mismo objeto o hecho del mundo físico, si bien desempeña un rol muy relevante en materia de verdad, no es lo único que cuenta. Esta premisa es importante para la tesis que defiendo: que aun cuando una prueba guarde relación causal con un hecho ilícito eso no la convierte en prueba ilícita si no está conectada desde el punto de vista normativo.

El concepto de regla constitutiva de Searle (1995) es utilizada para considerar que el derecho funciona estipulando que ciertos hechos brutos deben concebirse de cierta manera, según determinados contextos institucionales. Lo anterior me permite establecer la premisa que un hecho bruto se rige por relaciones de tipo causal, mientras los hechos institucionales, muy frecuentes en el derecho, están atravesadas por reglas que le asignan significado normativo y están regidos por relaciones de normatividad.

En la sección 5 resalto la ambigüedad del término prueba que es destacada por algunos teóricos de la prueba, para poner énfasis en que cuando hablamos de la prueba de los hechos, podemos estar haciendo referencia a problemas muy diversos.

En la sección 6 me valgo de la teoría de Davidson para sostener que una misma acción puede ser descripta de varias formas, por lo que aunque tengamos una única acción podemos tener diversas descripciones de dicha acción. También argumento que un mismo hecho puede dar lugar a pruebas lícitas o ilícitas según la descripción de la acción y según el contexto institucional en el que transcurra.

Todo lo anterior contribuye a fundamentar que la regla de exclusión probatoria en aplicación de la teoría del fruto del árbol envenenado, como la ha tratado la dogmática, la jurisprudencia y cierto sector de la teoría, trata a la prueba en términos causales o naturalísticos y no en términos normativos lo que, en ciertas ocasiones, ha contribuido a la confusión conceptual y a adoptar decisiones injustificadas, excluyendo pruebas que no deberían ser excluidas, o introduciendo excepciones a dicha regla sin una única línea argumental sólida que las respalde.

2. LA TENSIÓN ENTRE LA VERDAD Y LAS GARANTÍAS EN EL CASO DE LA PRUEBA TRIBUTARIA

En años recientes la dogmática, no solo tributaria sino de todas las ramas del derecho, ha analizado con detenimiento y atención cómo, desde el punto de vista jurídico, «prueba» obtenida por medios ilícitos provoca efectos jurídicos lícitos. El

desafío consiste en explicar las tensiones que se dan entre los distintos valores que los sistemas jurídicos democráticos persiguen como lo son la búsqueda de la verdad y el respeto de los derechos y garantías individuales fundamentales. La célebre afirmación del Tribunal Constitucional alemán: «la verdad no puede perseguirse a cualquier precio» (Tribunal Supremo Federal Alemán, sentencia de 14 de junio de 1960), da cuenta de dicho conflicto. Los sistemas jurídicos y los tribunales suelen aceptar que el derecho relegue la búsqueda de la verdad para la preservación de ciertos derechos fundamentales. Es imprescindible entonces pensar por qué razones debemos aceptar la postergación de ese valor de la verdad (si es que debemos hacerlo) y reflexionar si el conflicto o la tensión se presenta en los mejores términos analíticos y conceptuales por la dogmática contemporánea.

En la materia tributaria la prueba suele regirse por normas similares a las del resto de las disciplinas jurídicas y, desde el punto de vista epistémico, suele presentar los mismos desafíos (Ferrer Beltrán, 2005; Osorio Morales, 2023). La particularidad o los énfasis se podrían observar en que la relación jurídica tributaria y la obligación tributaria está trabada entre el Estado y el particular, de forma que de un lado de dicha relación se encuentran las Administraciones o Agencias tributarias que suelen tener ciertas potestades de investigación y fiscalización de los hechos generadores de tributos o hechos imposables con las que no cuentan los particulares. Esto implica que los Fiscos tengan mayores herramientas para investigar y que sea necesario que se acompañe de un profundo compromiso de respeto a los derechos fundamentales por parte de los funcionarios involucrados.

Dentro de dichas prerrogativas suelen estar la facultad de allanar domicilios particulares; inspeccionar domicilios fiscales, incautarse documentación, computadoras u ordenadores; requerir información a otros Estados en el marco de determinados convenios marco de cooperación para evitar el fraude internacional o la fuga masiva de capitales; citar a declarar a los contribuyentes, responsables y terceros a los efectos de recabar prueba; etc. Estas facultades especiales a la hora de diligenciar medios probatorios que poseen las Agencias tributarias, los sitúa, en buena medida, en una situación similar a la que tienen los policías cuando persiguen el delito, y por dicha razón, es muy relevante que actúen dentro del estricto marco de sus competencias. Esto impacta en que se pueda apreciar con claridad cuándo una prueba debe ser excluida por haber sido obtenida por medios ilícitos y en qué circunstancias una prueba podrá ser considerada ilícita o pruebas que deriven, directa o indirectamente de ella.

Desde el punto de vista estrictamente normativo, se suelen incorporar normas de presunción que facilitan la labor probatoria de las Administraciones fiscales porque suelen operar desplazando el objeto probatorio a otro que sea de más fácil prueba (suele ocurrir cuando se sustituye prueba de un hecho interno-una intención-, de mayor dificultad probatoria, por un hecho externo), o se establecen ficciones normativas que son normas fiscales *antielusivas* donde no hay una preocupación por los hechos del caso, sino una calificación jurídica del legislador que se antepone a cualquier valoración del juzgador de los hechos.

Todas las prerrogativas o facultades anteriores se justifican en la necesidad que tiene el Estado de recaudar ingresos para solventarse y cumplir con sus fines, que no son otros (o no deberían serlo) que garantizar los derechos individuales. Aunque, para determinada concepción, los tributos lesionan el derecho de propiedad de los individuos y la libertad no está presente en el marco de una relación jurídica tributaria, porque nadie elige pagar los impuestos que paga, no parece menos cierto que también garantiza la propiedad misma. Como han sostenido varios autores, sin tributos no es posible tener algo parecido al derecho de propiedad, ya que un Estado que garantice los derechos civiles y políticos mínimos (como el derecho a ser libre de la violencia de terceros y de gozar libremente de los bienes que se poseen) necesita contar con (en una visión mínima que parece poco discutible), una policía organizada que proteja a los ciudadanos del delito. Y sin recursos estatales esto no es posible (Holmes y Sunstein, 1999).

Por lo anterior, el interés general suele estar atrás de la recaudación de los Fiscos. Pero, en un Estado democrático, respetuoso de los derechos individuales y de la libertad personal, el interés general nunca puede conllevar el desconocimiento de estos derechos. Por dicha razón, es necesario ser cauteloso cuando se afirma que hay que hacer un balance de los distintos derechos involucrados o un ejercicio de ponderación, para no incurrir en el conocido aforismo «el fin justifica los medios», que pueda implicar el sacrificio de un derecho individual o de, lo que es lo mismo, una persona.

Reafirmando las consideraciones precedentes, pretendo analizar algunos casos jurisprudenciales europeos y latinoamericanos que se enfrentaron (al menos aparentemente) a esta situación dilemática de tener que sopesar entre la verdad y las garantías individuales. Concretamente, voy a tomar como ejemplo el impacto que tuvieron en la jurisprudencia de los Tribunales Constitucionales y Supremos europeos (particularmente en España), el célebre caso tributario *la lista Falciani* y un caso, también tributario, mucho menos conocido a nivel internacional, que es el caso de un frigorífico uruguayo donde se aplicó la regla de exclusión probatoria, en un sentido opuesto al pronunciamiento de los tribunales europeos. A nivel internacional, la regla de exclusión probatoria (que se originó en la jurisprudencia estadounidense) está siendo, cada vez más, objeto de excepciones y de deflación en su aplicación jurisprudencial (Miranda, 2010; Muñoz, 2023). El caso uruguayo es una excepción en esta corriente que se da en los Estados Unidos y en Europa.

3. LOS CASOS TRIBUTARIOS DILEMÁTICOS Y LAS REGLAS DE EXCLUSIÓN PROBATORIA

3.1. El caso Falciani

En el caso de la Sentencia Número 116/2017 de 23 de febrero de 2017 el Tribunal Supremo de España Sala de lo Penal (STS 116/2017, de 23 de febrero de 2017)

aceptó la validez de la prueba de la llamada «lista Falciani». Falciani (Belarmino era su nombre real), un empleado del banco suizo HSBC, había extraído ilícitamente información de datos bancarios de una serie de clientes del banco que residían en diversos países europeos. En el caso, Darío, un residente español, había sido objeto de reliquidación tributaria de los ejercicios 2005 y 2006 por no haber incluido en su declaración jurada de impuesto a la renta y patrimonio la información de las seis cuentas que poseía en el banco suizo. Tampoco tributó en Suiza por dichos fondos. Como surge del numeral Séptimo de los Antecedentes de hecho:

el acusado fue objeto de actuación inspectora por parte de la Agencia Española de Administración Tributaria (AEAT) con relación a los datos que sobre él constaban, correspondiente a los ejercicios 2005 y 2006, y citado de comparecencia hasta en once ocasiones, sin asistir personalmente a ninguna, sino por medio de representante (STS 116/2017, de 23 de febrero de 2017, p. 4).

Previamente, la Agencia Española de Administración Tributaria (AEAT) había desplegado un operativo con carácter general para que todos aquellos que debieran haber incluido los activos que figuraban en el banco suizo y no lo hubieran hecho ajustaran sus declaraciones juradas. Entre quienes no lo hicieron figura el sujeto de la sentencia que estamos analizando. La actuación de la Agencia Tributaria había comenzado con un oficio pidiendo información a la autoridad fiscal francesa para que le facilitara toda la información, como resulta del numeral primero de los antecedentes de hecho.

La sentencia reconoce la complejidad del caso y analiza los argumentos de la defensa de Darío que sostiene, básicamente, que la documentación en la que se basa la acusación de la AEAT, consiste en la prueba obtenida ilícitamente por Belarmino (Falciani), quien no sólo no tenía autorización para acceder a ella, sino que menos la tenía para difundirla y que si se condena al acusado, se lo estaría haciendo por prueba obtenida de forma ilícita, la que violaría su derecho a la intimidad, que alcanza también a la información económica. La defensa cita jurisprudencia francesa, donde se consideró la prueba como ilícita y también que los datos habían sido objeto de tratamiento porque no eran los registros del HSBC.

El tribunal reconoce en el numeral tercero de los Fundamentos de Derecho que la prueba obtenida con violación de los derechos fundamentales debe ser excluida de la valoración probatoria. Sin embargo, en los numerales siguientes, descarta que la prueba ilícita pueda tener los efectos contaminantes que la defensa le asigna. Repasa la jurisprudencia de los tribunales supremos de otros países europeos, donde todos rechazan dicho efecto contaminante (incluido el francés, con una variación en su jurisprudencia, acompañada de la sanción de nueva normativa para combatir el fraude fiscal a gran escala).

La sentencia recurrida se había basado en el principio de no indagación, en el sentido de que no le cabía al Estado Español controlar la legalidad de la actuación de otros Estados europeos. El Tribunal Supremo matiza el alcance del principio de no indagación, considerando que es fundamental analizar el caso concreto y que corresponde huir de interpretaciones rígidas. Recuerda su rechazo de una prueba

que consistía en la entrevista de dos policías españoles a un detenido en la prisión de Guantánamo a la que considera un inexplicable limbo jurídico.

Esta sentencia inaugura la tesis, a nivel del Tribunal Supremo, que la regla de exclusión probatoria contenida en el artículo 11 de la Ley Orgánica 6/1985, de 1 de julio, del Poder Judicial (en adelante, LOPJ), se centra en la figura del Estado en su actividad del acopio de pruebas y no de un particular. Por dicha razón, el tribunal se refiere en alguna ocasión a la lista Falciani como «fuente de pruebas» y no como la prueba misma. Sostiene el tribunal en el fundamento de derecho sexto que: «El particular que por propia iniciativa desborda el marco jurídico que define la legitimidad del acceso a datos bancarios... no lo hace en nombre del Estado» (STS 116/2017, de 23 de febrero de 2017, p. 11).

La sentencia limita la regla de exclusión probatoria y su efecto contagio a los casos donde quien comete la ilicitud es un agente del Estado, intentando obtener pruebas. Interpreta el artículo 11 mencionado como una norma que opera en el marco de un proceso penal, entendido en un sentido muy amplio, que regula la actividad de acopio de pruebas. Sostiene que su tesis no puede entenderse como una regla general, que el particular cuya actividad se dirige al acopio de pruebas para ser utilizadas en un proceso judicial también debería ser descartada, asimilándola a un agente del Estado y sostiene que no puede llegarse a la misma conclusión sin ponderar los hechos del caso y sin considerar si se trata de la violación del núcleo duro de un derecho fundamental o de su periferia (Fundamentos de Derecho séptimo y octavo).

La sentencia también analiza cuestiones vinculadas a la valoración de la prueba y su veracidad. La defensa había alegado que la única prueba de cargo era un archivo excel que estaba claro que había sido manipulado y que su autenticidad era dudosa, al mismo tiempo que sostenía la ruptura de la cadena de custodia de la prueba. Estos argumentos son desechados por el tribunal que asegura, en base a toda la prueba agregada, que no se ha visto afectada la incolumidad de la cadena de custodia de dicha prueba.

3.2. El caso del frigorífico uruguayo

Otro de los casos a considerar es el caso uruguayo, al otro lado del Atlántico, el que refiere también a un caso tributario. Se trata de la Sentencia Número 170/22 de 17 de marzo de 2022 del Tribunal de lo Contencioso Administrativo del Uruguay (STCA 170/22, de 17 de marzo de 2022) (único y supremo órgano en dicha fecha capaz de juzgar la legalidad de los actos administrativos del Estado uruguayo), la que, mediante mayoría de tres miembros contra dos, considera que se violaron las garantías del debido proceso por aceptar una prueba obtenida por medios ilícitos y se decanta por aplicar la teoría del *fruto del árbol envenenado*, considerando nula toda la prueba que constaba en el expediente.

Un expleado de la empresa (el Sr. Manuel) que había sido despedido por notoria mala conducta por cometer un delito continuado de estafa contra la empresa, se presentó en las oficinas de la Administración Tributaria (Dirección General Impositiva) para realizar una denuncia contra el frigorífico, para el cual se había desempeñado como jefe de administración. Junto con la denuncia presentó un DVD, una planilla con claves, un disco portable y fotocopias encuadradas. En su denuncia sostenía que la empresa no facturaba todo lo que vendía, sino que utilizaba un sistema informático que le permitía facturar el 25 por 100 o el 50 por 100. Para hacerlo, dicho sistema permitía que la empresa llevase una doble contabilidad o doble registración mediante la cual se mantenían separados los registros: por un lado, en lo que se denominaba CUBO PRU, se registraban todas las operaciones y en el CUBO FC se registraban las operaciones que se facturaban. El sistema tenía una operatividad de 24 horas que permitía que los registros se borrarán sin que quedase rastro de la maniobra.

Surge de los antecedentes reseñados en el Considerando II de la Sentencia que las inspectoras originalmente recibieron la denuncia y la documentación. Posteriormente, al no poder abrir los discos entregados por el denunciante en los ordenadores de la Administración Tributaria, citaron al denunciante para que compareciera. El denunciante compareció con su propia computadora y abrió el DVD con los archivos informáticos. La DGI no tenía un software que le permitiera abrir los documentos porque se trataba de un software a medida hecho para la empresa por un proveedor informático. El denunciante explica la funcionalidad del sistema y relata por qué hay ciertos comprobantes que tienen ciertas inscripciones y marcas y explica su significado.

Del considerando III surge que las inspectoras labraron un acta a la empresa donde esta le entregó 384 comprobantes que retiraron de sus registros informáticos y tomaron declaración a la directora de la empresa y a dos administrativos que se desempeñaban en ella. También que fueron al proveedor del software y que éste le proporcionó copias de correos electrónicos de donde surgía que la empresa le había solicitado un programa hecho a medida con determinadas funcionalidades. También se tomó declaración a los socios de la empresa proveedora del software que permitía dividir las ventas, una con IVA y otra sin IVA.

En el considerando (fundamento de derecho) IV.1, el tribunal aclara: «En primer lugar, cabe analizar el agravio referido a la ilicitud de la prueba cuestionada que, según los accionantes, fue la base del acto enjuiciado» (p. 16) El tribunal aclara que no resulta controvertido que la Administración tributaria tuvo conocimiento de la doble contabilidad y del sistema informático diseñado, eventualmente, para defraudar tributos gracias a la prueba aportada por el Sr. Manuel. La sentencia cita uno de los votos de los ministros, donde se alega una cuestión diferente, pero vinculada con la ilicitud de la prueba. Allí se sostiene que la documentación en base a la cual la Administración hace su reliquidación de impuestos proviene de la prueba clandestina aportada por el denunciante. De esta forma, al no constar la veracidad de dicha prue-

ba, y al no seguirse, en opinión de la mayoría, las reglas para la copia de la documentación, la prueba así obtenida es nula porque viola las garantías del debido proceso. Se sostiene que la documentación debió extraerse de la computadora del frigorífico y ser copiada en presencia de los representantes de dicha empresa. Posteriormente se afirma que no solo es ilícita la prueba obtenida por medios ilícitos sino la refleja que derive de ella, aunque sea en sí misma considerada lícita, como lo es la practicada por la Administración Tributaria.

Conviene repasar que la mayoría del tribunal, aunque los trata en forma conjunta, se basa en dos cuestiones conceptualmente separables: la ilicitud de la prueba agregada por el denunciante y la violación del debido proceso al no asegurarse la cadena de custodia probatoria. Sin embargo, esta segunda cuestión parece ser objeto de un malentendido, dado que de los propios antecedentes de hecho señalados por la misma mayoría, surge que la propia empresa entregó los 384 comprobantes de ventas, que son facturas oficiales de las ventas de la empresa que, a su vez, son documentos emitidos por la propia empresa y entregados a la Administración tributaria. Esta errada valoración de los hechos probados será considerada más adelante en el punto 6.2.

La discordia sostiene que la prueba ilícita proporcionada por el denunciante no contamina el resto de la prueba obtenida porque la Administración recibió la denuncia y comenzó una investigación de fuente independiente. Especialmente, la investigación realizada al proveedor del software donde surgió con total claridad de los correos electrónicos y de las declaraciones de los socios de la empresa que se le había solicitado un programa que permitiera partir la facturación en dos partes. La discordia considera que la «exclusionary rule» admite diversas excepciones como son el vínculo atenuado, la fuente independiente y la inevitabilidad del descubrimiento. No hace referencia alguna a que la prueba haya sido obtenida sin respetar la llamada incolumidad de la cadena de custodia y tampoco confuta este argumento sostenido por la mayoría. Pone énfasis sí en que la prueba fue lícitamente obtenida por la Administración mediante causas de fuente independiente y que, por lo tanto, la prueba no fue ilícita.

4. EL PROBLEMÁTICO CONCEPTO DE HECHO EN EL DERECHO

En el discurso cotidiano solemos usar la palabra hecho con significados muy diversos y en contextos muy dispares. Resulta imprescindible que cualquier teoría acerca de la prueba, se implicara con esta problemática porque la prueba, es precisamente la prueba de los hechos.

Desde Aristóteles (1994), la noción de hecho viene muy emparentada con la trilogía de creencia, verdad y justificación. El dicho popular «nada más obstinado que los hechos», reafirma esta idea. Tenemos razones para creer en los hechos. La razón

teórica, de la que estamos dotados en tanto seres humanos racionales, nos da razones para creer. La razón práctica nos da razones para la acción. Las razones para la acción, presupone (en muchos casos) razones para creer. Los jueces son agentes racionales que están llamados, por los sistemas normativos en los que se desempeñan, a tomar decisiones. No hay nadie que afirme que sea algo deseable o bueno tomar decisiones con una información deficitaria o falsa. Todos consideramos que contar con más y mejor información nos lleva a tomar mejores decisiones. Los jueces están especialmente motivados por conocer los hechos. Cualquier retaceo a esta necesidad material y lógica debe estar debidamente justificada y corre con la carga de la argumentación.

Ahora bien, ¿qué significan los hechos en el mundo del derecho? Para un científico en un laboratorio los hechos son el cambio molecular, el comportamiento de la bacteria x, del virus z, etc. Estos hechos los llevan a tomar determinadas decisiones, por ejemplo, realizar una prueba con determinada sustancia química. Para una maestra escolar el hecho que le lleva a cambiar su metodología didáctica en el dictado de sus clases es el comportamiento de sus estudiantes en el aula, por ejemplo, que suelen distraerse, conversar entre sí, etc. Los hechos en el mundo jurídico no son desnudos, o brutos, o no se comportan, por decirlo de alguna manera, de una forma *pura*.

Antes de explicar esto último de forma más detenida, volvamos a Aristóteles. Cuando Aristóteles presentó su célebre fórmula de la verdad como correspondencia, en esa frase que después hizo también célebre y contemporánea Tarski (1969, p. 4), «la nieve es blanca “si y solo si”, la nieve es blanca», Aristóteles deploraba una verdad como consenso y sostenía que el enunciado «la nieve es blanca» no es verdadero porque todos afirmemos que la nieve es blanca, sino por el *hecho* de que es blanca. Para entender mejor esta afirmación de Aristóteles tuvimos que esperar unos siglos más tarde (y no significa que la hayamos comprendido a cabalidad) a Bertrand Russell (2013) quien sostuvo que, un hecho es aquello que hace verdadera a una proposición.

Los teóricos del derecho y la prueba han sostenido que en el derecho no trabajamos sobre los hechos, sino sobre enunciados acerca de los hechos (Gascón Abellán, 2010; Ferrer Beltrán, 2007; Taruffo, 2008). Lo anterior, podemos afirmar, tiene un doble significado. En primer lugar, uno bastante obvio: que los juristas trabajamos con lenguaje, con expedientes judiciales, con documentos. En un segundo sentido, bastante menos obvio y menos enfatizado: que los hechos en el derecho, además de estar mediados lingüísticamente, están atravesados por reglas que, en algún sentido, los crean.

En el primer sentido afirmado (el que trabajamos con lenguaje), vale destacar que esto implica que no consta agregada al expediente el arma homicida, el cadáver, el veneno, la cuenta bancaria, etc. Tenemos representaciones de los hechos, que operan como una suerte de «doble». Tenemos el informe del médico forense perito que analizó el cadáver; el informe pericial del arma homicida; el informe toxicológico; el informe del perito contable-financiero; la declaración de la testigo sobre lo que observó. En definitiva, *pedazos* de lenguaje que pretenden ser representaciones fieles de pedazos del mundo. El lenguaje, en su función descriptiva o asertiva, afirma un

enunciado cuya verdad representa un hecho del mundo. También sabemos que no existen sólo los hechos externos, sino también los hechos internos, que son los estados mentales de un individuo y que implican que sea o no verdadera la afirmación de que *tenía la intención de cometer fraude tributario o de ocultar su patrimonio del Fisco*. También hay hechos voluntarios e involuntarios. Los hechos voluntarios involucran la conducta humana. Hay hechos de corte positivo o afirmativo, y los hay de corte negativo, como son las omisiones.

En el segundo sentido afirmado (el de las reglas constitutivas o institucionales), el derecho es, de alguna forma, una máquina de crear *hechos*. Hacemos bien si esta afirmación enciende nuestras alarmas. ¿Qué significa que el derecho crea hechos?

4.1. Los hechos y las reglas constitutivas o institucionales

La afirmación de que el derecho crea hechos no debe entenderse como apoyando corrientes escépticas en materia de prueba. No afirma que el derecho se desinterese por la verdad genuina de los hechos porque tiene reglas técnicas o de otro tipo que le permite desentenderse de la verdad en procura de otros valores que no tienen que ver con ella. Asumiendo una corriente *objetivista crítica no ingenua* en materia de prueba (González Lagier, 2005), podemos afirmar que el derecho, de alguna forma, en un sentido relevante, crea hechos.

Le debemos a la corriente de la filosofía del lenguaje *ordinario* haber reparado en que hacemos cosas con palabras (lo que da cuenta de otra dimensión del vínculo entre mundo y lenguaje). Esta afirmación, que se debe a la teoría de John L. Austin (1991), puso en claro que el lenguaje no solo posee una dimensión descriptiva, junto con una prescriptiva, sino que también posee una dimensión *performativa* o pragmática. Un enunciado en función descriptiva del tipo «hoy el día está soleado» pretende describir el mundo y que el enunciado sea verdadero va a depender efectivamente de que el día esté soleado. Las normas jurídicas suelen estar formuladas en enunciados de tipo prescriptivo, es decir, prescriben la conducta que indican. El enunciado «cierra la puerta» no es verdadero o falso. Aunque, por supuesto, parte de ciertas condiciones de carácter material que tienen que estar dadas para que el enunciado sea un enunciado con sentido (que haya una puerta y que esté dirigida a un interlocutor que tenga la capacidad de cerrarla).

La función pragmática del lenguaje da cuenta de que con las palabras hacemos cosas y cambia el «estado del mundo». Así, el insulto «te comportas como un troglodita» no pretende describir el estado del mundo, sino proferir un insulto que incomode al interlocutor. Si mediante una expresión como «buen día», nos damos por saludados, mediante un insulto seguramente nos sentiremos ofendidos y provoquemos la aparición de un hecho interno, nuestro estado de ánimo ofendido, que antes no existía en nosotros. Este hecho interno causado por el insulto proferido por nuestro interlocutor puede ser a su vez una causa de un nuevo hecho externo, etc.

Cuando una pareja presta su consentimiento frente a determinada autoridad oficial para contraer matrimonio, por la sola afirmación del tipo «acepto», cambia su posición en el mundo, su estado civil, y eso implica que tiene derechos y obligaciones que antes no tenía; por haber pronunciado determinadas palabras en determinado contexto.

Continuando con la línea de Austin, John Searle, en su clásico *Actos de habla* (2017), desarrolló la idea de que es necesario distinguir las reglas regulativas de las reglas constitutivas. Las reglas regulativas regulan conductas que ya existen en el mundo. Una regla de tránsito que establezca que los vehículos deben conducir a una velocidad máxima de 60 km por hora, parte de hechos que existen con independencia de las reglas de tránsito, estos hechos son la existencia de máquinas (vehículos, objetos del mundo) y la posibilidad de que se desplacen a cierta velocidad.

Las reglas constitutivas son propias de los contextos institucionales, pueden definirse como «x cuenta como y en el contexto c» y son aquellas que crean conductas que no existen con independencia de las reglas. Un ejemplo paradigmático son las reglas de los juegos. Consideremos, por ejemplo, la conducta hacer «penal» en el fútbol, la norma que sanciona el penal, también lo define. Cualquier falta (una patada, un codazo) en el área realizada por un jugador del equipo contrario, o tocar la pelota con la mano, con excepción del arquero, implica cometer penal y que sea sancionada con un tiro directo, sin barrera, con una distancia de once metros. No existe la conducta «cometer penal» por fuera de la regla que sanciona el penal. Fuera de la cancha, lo que alguien hace es dar una patada o un codazo, no hacer un penal, aunque los hechos brutos, las externalidades, sean las mismas. Si un legislador levanta la mano en un recinto parlamentario, acaba de sancionar una ley. Cuando sale del recinto parlamentario y camina unos pocos metros, y hace exactamente lo mismo, simplemente detiene un taxi.

Por supuesto que, en estos ejemplos, la finalidad de la conducta (cierto aspecto de los hechos internos) podríamos pensar, son bien diferentes. El jugador comete una sanción para detener al jugador contrario (en la mayoría de los casos), el parlamentario quiere sancionar una ley, no parar el taxi, o no desperezarse. De todas formas, en todos estos casos, la conducta también es intencional. Desde este punto de vista, los hechos internos también coinciden parcialmente: no son negligencias, o conductas culposas; son conductas todas intencionales. Los estándares del derecho que son la culpa o el dolo, no nos sirven tampoco para diferenciar estas conductas, que tienen un componente también de hecho interno que coincide. Ahora bien, ¿qué sucede con la finalidad de la conducta? Hablamos del fin que persigue la conducta en los diferentes contextos en los que se desarrolla. ¿Podría perseguirse esta finalidad, este objetivo diferente, sin la existencia de las reglas institucionales? Parecería que no, que el agente sólo puede perseguir esta finalidad sirviéndose de las reglas institucionales que existen. Su finalidad, su ámbito estrictamente subjetivo del fin deseado, no puede ser perseguido sin que operen las reglas institucionales. Es decir, no compren-

demostremos la conducta subjetiva si no apreciamos, al mismo tiempo, la existencia de las reglas objetivas.

El derecho no recoge hechos brutos, hechos desnudos, regularidades de comportamiento exclusivamente, califica los hechos, los atraviesa por las reglas y los resignifica. En el mundo compuesto por las regularidades de comportamiento, existe la conducta «utilizar un documento falso», podríamos decir, pero no la conducta cometer fraude o defraudación tributaria. Uno no observa en el mundo fraudes tributarios. Uno puede observar que Juan le pidió a Pedro que le extendiera una factura que no recoge la verificación de ningún hecho generador o hecho imponible tributario, esto es, no le enajenó ningún bien, no le prestó ningún servicio. El documento no recoge la realidad. Nosotros decimos que la conducta, el hecho, fue cometer fraude tributario. Pero, reitero, la conducta «cometer fraude tributario» no existe en el mundo observacional. Con esto no quiero decir que se fomente la arbitrariedad de las Administraciones o Agencias tributarias, quiero decir que tenemos que tener precaución para conocer el alcance del concepto de hecho. Porque la prueba es, o debería ser, la prueba de los hechos.

En el clásico modelo subsuntivo, donde se trata de subsumir o encajar la premisa fáctica en la premisa normativa, nuestra premisa fáctica ya está, en buena medida, atravesada por la calificación jurídica, por los llamados problemas de calificación. Si pensamos en el siguiente ejemplo simplificado de silogismo *modus ponens*: «El que cometa fraude tributario deberá ser sancionado con la multa o la pena de...»; «Juan cometió fraude tributario», entonces o, en conclusión, «Juan debe ser sancionado con la pena de...», la premisa menor o premisa fáctica del razonamiento, además de los problemas de prueba, tiene los problemas de calificación, que se agregan a las clásicas restricciones de corte epistémico (Moreso y Vilajosana, 2004).

Lo anterior quiere decir que podemos saber perfectamente qué significa en un determinado sistema jurídico cometer defraudación o fraude tributario, podemos saber o tener muy claro, cuáles son los hechos del caso, en el sentido de regularidades de comportamiento, saber qué paso (sin lugar a la menor duda) en el mundo observable, pero todavía no podemos saber si Juan cometió defraudación tributaria; este es un problema de calificación. No sabemos si los hechos de este caso, si los hechos particulares a los que nos vemos enfrentados, son o no, un caso de *instanciación* del hecho general descrito en la premisa normativa o premisa mayor de nuestro razonamiento¹.

Por ejemplo, si Juan cuando realiza una enajenación de un bien no extiende una factura o un comprobante fiscal y lo hace deliberadamente, todavía no podemos

¹ Este problema se debe, en buena medida, a la vaguedad de ciertos términos utilizados por el derecho en general. No obstante, no es el objeto de este artículo problematizar en torno a esta compleja cuestión. Basta señalar que la conceptualización de cuándo nos encontramos frente a un caso de evasión o defraudación tributaria o frente a un caso de elusión tributaria depende de la aceptación de determinados estándares valorativos por parte de la comunidad de operadores jurídicos.

saber si Juan cometió o no fraude tributario. La prueba de los hechos y los hechos son estos: que hubo una venta en su establecimiento comercial y que esa venta no la registró. Muchos sistemas jurídicos tributarios, como dije al comienzo, establecen normas legales de presunción y consideran que la no extensión del documento requerido es una presunción del dolo defraudador. Esto se debe a que es muy difícil probar el ánimo de alguien, porque la prueba de un hecho interno directa es casi inaccesible, entonces la ley, frecuentemente, establece normas de presunción que permite al juzgador considerar probado un hecho, si antes prueba la verificación de un hecho diferente. Así funcionan las normas de presunción, funcionan con la fórmula «Dado P debe procederse como si Q fuese verdadero» (recientemente Mendonca (2022) ha calificado las normas de presunción como reglas constitutivas). Sin embargo, en mi opinión, eso no basta para concluir que Juan cometió defraudación. Porque, aunque su conducta de no extender la factura sea intencional, puede suceder que, aun así, no haya verificado defraudación tributaria. Puede haberlo hecho por otra motivación que no es la de pagar menos impuestos al Fisco, sino porque tiene algún problema técnico en el sistema de facturación, o simplemente había mucho público en el local comercial, o se trata de un establecimiento que no tiene instalado el sistema y posteriormente va a extender la factura o va a pagar los impuestos que correspondan. Pero estas últimas valoraciones o hipótesis que compiten con la del ánimo defraudador, involucran problemas de calificación, aunque también involucren problemas de prueba. Podemos discutir, desde el punto de vista de la calificación, si el sólo hecho de no extender la factura computa o no como el hecho de cometer *defraudación o fraude tributario*.

En lo que aquí importa, hablar del hecho de que Falciani copió una base de datos que contenía información reservada de los clientes del banco HSBC, no es un hecho bruto, entendido como regularidad de comportamiento, como una acción intencional como quien toma un vaso de agua para calmar su sed, porque «base de datos», «reservado», «información», son hechos atravesados por determinadas reglas constitutivas o lo que la dogmática suele mencionar como las reglas aplicables al caso.

Cuando Manuel (el empleado supuestamente «infiel» del frigorífico) respaldó durante años, con detalle y esmero todos los archivos informáticos de la empresa que daba cuenta de los comprobantes que se fraccionaban para defraudar al fisco uruguayo, el hecho bruto, la regularidad de comportamiento es única: copiar los archivos. Pero *normativamente* ese hecho, según las reglas que definamos y en el contexto institucional que nos movamos, pueden ser significativamente cosas bien diferentes: la prueba de su participación en el delito tributario; la prueba que lo exculpe si se lo acusa de estafa (delito por el que resultó, finalmente, condenado); la prueba, en parte, de su estafa.

Esto nos sirve para observar que la misma regularidad de comportamiento, con la misma intención, puede significar *normativamente* pruebas o hechos muy diferentes. Por dicha razón varios autores han remarcado la ambigüedad con la que los juristas hablamos de hecho y prueba (Ferrer Beltrán, 2005).

Como dice Searle (1995), los hechos institucionales existen porque las personas podemos asignarles funciones a las cosas: así la piedra puede ser un pisapapeles y la luna puede ser algo bello.

4.2. Los hechos y los niveles de lenguaje

Vimos en los puntos anteriores que los hechos y la verdad son conceptos muy entrelazados. No existen los hechos falsos, no puedo conocer un hecho falso, para que algo sea un hecho, tiene que ser, necesariamente, verdadero. Aunque, estrictamente hablando, un hecho es un hecho, y lo único que, desde el punto de vista conceptual, tiene un valor de verdad o un valor *veritativo* (verdadero o falso), son los enunciados acerca de los hechos y no los hechos mismos. Los hechos tornan verdaderos los enunciados, cuando no tenemos un hecho que respalde un enunciado, entonces el enunciado es falso.

Lo anterior, no es así en todos los casos, o en todos los casos donde proferimos un enunciado; es necesario saber en qué contexto estoy hablando o profiriendo el enunciado en cuestión o, mejor dicho, cuál es mi lenguaje objeto. Por ejemplo, el enunciado «Batman es el héroe de ciudad gótica» o «el mejor amigo del Quijote es Sancho Panza», ese enunciado es verdadero, porque se ajusta al mundo con el cual es necesario contrastar el enunciado. En las películas o las tiras de cómics efectivamente es verdad que Batman es el héroe de Ciudad Gótica. En la novela «El ingenioso hidalgo Don Quijote de la Mancha» es verdad que Sancho Panza es el mejor amigo de Don Quijote.

El enunciado «Batman es el héroe de Ciudad Gótica» está formulado en un nivel 1 de lenguaje o metalenguaje, que se corresponde con el lenguaje objeto que es la película de Tim Burton. Popper (2003) dijo que hasta que Tarski presentó su trabajo, él no se animaba a hablar de verdad como correspondencia porque miraba el concepto con desconfianza. La gran virtud de Tarski fue presentar su teoría de la verdad siempre en términos de un lenguaje objeto y un metalenguaje. Un enunciado es verdadero en un metalenguaje cuando se corresponde con otro enunciado de un lenguaje objeto. El enunciado proferido por el tribunal constitucional español «la Agencia tributaria española concluyó que Darío no declaró sus cuentas al Fisco español» (STS 116/2017, de 27 de febrero de 2017, p. 3), es verdadero si y solo si «la Agencia Tributaria española concluyó que Darío no declaró sus cuentas al Fisco español». Por dicha razón, la teoría de Tarski (1969) es conocida como una teoría de la verdad *desentrecomilladora*, porque lo que se encargó de establecer es una teoría de la verdad para sistemas formalizados donde se conceptualizase con claridad la distinción de distintos niveles de lenguaje. La teoría de Tarski no presentó una teoría de la verdad donde a una determinada porción de la realidad o de hechos del mundo objetivo, se correspondiera determinada porción de lenguaje. Su teoría es sobre distintos niveles de lenguaje.

En ese sentido, los hechos que usa o emplea el razonamiento probatorio del agente racional juzgador, son los enunciados acerca de los hechos que, a su vez, están enunciados también en lenguaje objeto. Solemos tener entonces tres niveles de lenguaje: el lenguaje objeto, el metalenguaje 1 y el metalenguaje 2. En el lenguaje objeto más mínimo podríamos tener algo así como ciertos hechos brutos, o desnudos. El hecho de que alguien colocó su patrimonio en cuentas bancarias en otro país y que no denunció el contenido patrimonial de dichas cuentas o activos a las Agencias tributarias. El hecho «cuenta a declarar», «operación a declarar», «activo a declarar», es, rigurosamente hablando, algo más que un simple hecho bruto, porque está atravesado por las reglas constitutivas que analizamos en el apartado anterior e involucra problemas de calificación, pero ignoremos esta situación a los efectos del argumento. En el lenguaje objeto sucedieron estos eventos: Darío no declaró sus cuentas al Fisco español. En un metalenguaje, o nivel 1 de lenguaje un informe o la denuncia de la Agencia tributaria española nos dice que: «Darío no declaró sus cuentas al Fisco español». En un nivel de metalenguaje 2, el juzgador, típicamente el juez, está analizando la verdad entre el lenguaje objeto, o los hechos del «mundo» y el metalenguaje 1, está analizando si hay correspondencia entre los dos niveles de lenguaje.

Lo anterior puede complejizarse mucho más y se podrían distinguir muchos niveles de lenguaje. Así, las Administraciones o Agencias tributarias son organismos complejos que cuentan con numerosos órganos que la componen que van tomando decisiones y se la traslada a otros órganos de la estructura jerárquica, de forma que, por ejemplo, un inspector de hacienda constata determinada situación mediante un acta, generalmente hace un informe que lo eleva a su superior que, a su vez, hace otro informe y después se eleva a una asesoría letrada que también hace otro informe, que, a su vez, lo eleva el jerarca máximo del organismo para que dicte una resolución o determinado acto administrativo. Aquí tenemos el lenguaje objeto o los hechos del mundo, que son que alguien tiene una cuenta que no declaró, el acta del funcionario que es el metalenguaje 1 de estos hechos, el acta de su superior que es el metalenguaje 2 y así podemos seguir. Pero también, el acta del funcionario es el lenguaje objeto de su superior. Quiere decir que podemos conceptualizar algo como lenguaje objeto o metalenguaje según el nivel de lenguaje en el que nos encontremos.

Lo anterior adquiere relevancia en el terreno de la prueba, porque si los hechos están compartimentados en distintos niveles de lenguaje, es necesario concluir que tenemos diferentes niveles de la prueba también, y diferentes niveles de reglas operando en cada tramo de los hechos o niveles de lenguaje respectivo. Por ejemplo, la prueba de los hechos del inspector de hacienda es el informe del banco y la declaración de Darío de que no declaró sus cuentas al Fisco, hipotéticamente (como vamos a ver más adelante en el caso que nos ocupa de «la lista Falciani», Darío nunca se presentó a declarar a las oficinas de la Agencia Tributaria española, lo que deviene relevante). El superior jerárquico tiene como objeto o prueba de su informe, el informe del inspector de Hacienda y, aunque desde el punto de vista jurídico, no es técnicamente lo que la dogmática conceptualiza como prueba (sino la formación

de la voluntad administrativa del órgano estatal), vamos a ver que no deja de ser un aspecto importante en el ámbito de la prueba de los hechos. Y lo es cuando se involucran reglas de exclusión probatoria, específicamente, porque es necesario conceptualizar qué queremos decir cuando hablamos de la prueba de los hechos, cuáles son las pruebas y en qué sentido están conectadas causal o normativamente.

El problema de los hechos también se ve cuando se analiza la prueba pericial. Por ejemplo, el juez no analiza el cadáver él mismo, los hechos brutos del caso, analiza el informe del médico o perito forense. Ese es el lenguaje objeto del juzgador y es, en conclusión, en base a lo cual construye su premisa fáctica. Es la forma que tiene el juzgador de acceder a los hechos del caso. Él no puede conocer directamente los hechos del caso porque experimenta restricciones epistémicas que el perito, en tanto técnico experto, viene a subsanar. La teoría de la prueba ha analizado en detalle los problemas que enfrenta la prueba pericial (Vázquez, 2013). Estos problemas, que podría resumirse en los problemas que enfrenta «el lenguaje de los hechos» (nótese lo paradójico de la afirmación porque se supone que los hechos no son lenguaje), repercute en la prueba y su impacto se observa también en los dilemas que atraviesan las reglas de exclusión probatoria, como vamos a tener ocasión de analizar más adelante.

5. LA AMBIGÜEDAD EN EL TÉRMINO PRUEBA Y PRUEBA ILÍCITA

Tanto en el caso de *la lista Falciani*, como en el caso del frigorífico, se observa una ambigüedad en el término de prueba. Como ha señalado recientemente Ho (2021), en un primer sentido, los juristas denominan «prueba» a lo que los epistemólogos consideran como «los objetos de la evidencia sensorial». En esta categoría es posible englobar la prueba oral; la prueba documental y la prueba «real» (el arma homicida, por ejemplo). En un segundo sentido, se habla de prueba también para referirse al enunciado, al lenguaje, que es una proposición que se establece sobre la prueba en su primera acepción. Por ejemplo, se afirma que el hecho de que el acusado estuviera próximo a la escena del crimen (prueba en el primer sentido) es prueba o un indicio probatorio (en el segundo sentido) de su participación en el homicidio. En un tercer sentido, se habla de prueba como premisa fáctica que es relevante por su rol inferencial. Nos permite inferir algo relevante para el corazón del caso. Por ejemplo, las huellas dactilares del acusado encontradas en la escena del crimen permiten inferir que él estuvo ahí y que fue quien asesinó a la víctima. En este tercer sentido el concepto de prueba es inferencial. En un cuarto sentido, algo puede tomarse como prueba en la vida cotidiana (rumores) pero no ser admitido como prueba en el derecho. Esta es la cuestión que involucra el concepto de «prueba lícita». Por lo que, en este cuarto sentido, por prueba puede entenderse las condiciones de admisibilidad de dicha prueba.

El concepto de prueba tiene también la clásica ambigüedad proceso/producto. Cuando, por ejemplo, se habla de las reglas de la prueba, se hace referencia al proceso de obtención de las pruebas. Cuando se habla de prueba ilícita se hace referencia al resultado o producto del proceso de obtención de las pruebas. La diferencia de la

terminología se encuentra también en las teorías que problematizan en torno a la cuestión de la relación de la prueba y la verdad. La relación entre verdad y prueba es conceptual en la primera tesis. Para que un enunciado esté probado, la verdad del enunciado es condición necesaria aunque no suficiente. Para la otra tesis, la conexión entre verdad y prueba es teleológica, la verdad es el fin a conseguir mediante la actividad probatoria. La primera tesis refiere a la prueba como resultado y la segunda como procedimiento (Ferrer Beltrán, 2005).

Gascón Abellán (2010) también señala que, en los contextos de descubrimiento y los contextos de justificación, el uso del concepto de prueba es bien diferente. En el marco del contexto de descubrimiento *prueba* hace referencia al procedimiento mediante el cual se llega a la formulación de enunciados asertivos verdaderos. En el marco del *contexto de justificación* se trata de las razones (prueba) que permiten considerar estos enunciados como justificados.

Podríamos seguir dando ejemplos de la polisemia que padece el término prueba y su uso en contextos muy diferentes, sin embargo, eso excede ampliamente el objetivo del presente trabajo. En lo que aquí nos importa, buena parte de la dogmática y la jurisprudencia no ha observado la ambigüedad del término prueba ilícita. Cuando se trata de condenar al culpable y usar la prueba en su contra, nadie podría justificadamente sostener que la prueba es ilícita, porque es precisamente lo que la dogmática llama «el cuerpo del delito». La lista Falciani, el archivo electrónico conteniendo la información hurtada al banco es la prueba del delito cometido por Belarmino (Falciani). Sólo cuando la prueba es utilizada para otro procedimiento distinto tiene sentido preguntarse si es ilícita. Podría serlo en un proceso laboral o civil donde Belarmino fuese el demandante y quisiera servirse de dicha prueba para obtener un beneficio económico. En ese caso sería aplicable el principio general de derecho «nadie puede beneficiarse de su propio fraude» y en dichos casos, cabe concluir, la prueba puede considerarse ilícita.

Idéntica situación se da en el caso del frigorífico uruguayo. Si Manuel es sometido a un proceso penal por haber hurtado información (proceso al que no fue sometido), los archivos informáticos sustraídos serán prueba fundamental de la eventual comisión del ilícito. En este sentido, la prueba y su licitud tienen que ser analizadas bajo parámetros divergentes.

6. LA DIFERENCIA ENTRE LA CAUSALIDAD Y LA NORMATIVIDAD Y LA REGLA DE EXCLUSIÓN PROBATORIA

La causalidad es una relación que suele darse entre fenómenos naturales: que empuje un determinado objeto es la causa de que el objeto se desplace. Que el fruto madure en el árbol es la causa de que caiga. El mundo natural se rige por la regla de la causa y el efecto. Los fenómenos naturales se pueden describir y «son». También pueden ser objeto de predicción. Se dice que los fenómenos naturales se explican;

pueden explicarse en virtud del mundo físico. Los fenómenos causales suelen involucrar cuestiones de hecho y las actitudes proposicionales de los sujetos acerca de estos hechos son creencias. En relación con los fenómenos causales, juego un rol relevante la razón teórica que me dice lo que tengo razones para creer, como vimos en la sección 4. Hume (2000) fue el primero en sostener que el hecho de que algo *sea* de determinada manera no significa que *deba* serlo. Hume se detuvo en esto considerándolo como una falacia².

El *ser* y el *deber ser* corren por carriles separados y la razón práctica involucra cuestiones de razones para la acción y nos interpela acerca de qué debemos hacer. El lenguaje en términos de *deber ser* involucra obligaciones y el vínculo es no ya de causalidad sino de *normatividad*. El problema de los hechos en el derecho plantea el desafío que se trata de describir y conocer los hechos, pero la construcción de la premisa fáctica es la de una premisa que se inserta en un razonamiento práctico. El agente que resuelve el caso, el juez, está encaminado a tomar una decisión. No aspira simplemente a conocer los hechos, como el científico en el laboratorio, el físico, el sociólogo o el antropólogo.

Desde el punto de vista normativo, el razonamiento práctico del juzgador involucra cuestiones de deber. Se plantea si se *debe* condenar al acusado. Es un razonamiento que tiene que ver con razones para la acción del juzgador. La premisa fáctica (los hechos del caso) desempeñan un papel instrumental y no por esto deja de ser de primera importancia, pero deben considerarse atravesados por el marco de este tipo de razonamiento práctico. Por supuesto que, si se razona en base a una premisa falsa, el razonamiento será defectuoso, incluso en tanto razonamiento práctico.

No es extraño que visiones escépticas a nivel de *metateoría* del derecho se valgan de este tipo de apreciaciones *naturalizadas* para negar la importancia de la verdad en el derecho. Por ejemplo, el *realismo* o *naturalismo filosófico* norteamericano, cuyo representante más conspicuo en la actualidad es el destacado filósofo del derecho estadounidense Brian Leiter (2007), sostiene un marco filosófico *naturalizado* que implica afirmar que el genuino conocimiento viene dado por las ciencias de la naturaleza o ciencias duras. Mientras que en el ámbito normativo (entendido en sentido amplio como abarcando no sólo al derecho sino también a la moral), la objetividad y la verdad no desempeñan ningún rol relevante. Esta tesis se afianza en sostener que el único conocimiento verdadero es el de los hechos porque los hechos admiten una posibilidad de control epistémico que los enunciados de deber bloquean, ya que no existen algo así como los hechos normativos o los hechos morales. Sin embargo, en mi opinión, los hechos no son imprescindibles para preservar la noción de verdad y de objetividad, porque las razones también nos permiten hacer un control de racionalidad. Y es posible preservar la tesis de la objetividad en materia normativa, preservando la posibilidad del error *normativo* (Korsgaard, 1996 y Street, 2011).

² Mucho después, Moore (1903) bautizó esta falacia como la falacia naturalista.

En el caso puntual de la regla de exclusión probatoria, tratándose de un problema que involucra la ilicitud de la prueba (haber sido obtenida en contra de determinadas normas jurídicas) llama poderosamente la atención que la dogmática ponga el foco en la cuestión causal. Esto se observa claramente cuando se piensa en el efecto contagio de la prueba ilícita, que precisamente suele nombrarse como «efecto». Lo que se analiza es si la prueba está conectada causalmente con la prueba originalmente ilícita. Pero tratándose de «ilicitud» tendría mucho más sentido que se conceptualizara en términos de imputación a los efectos de la conexión, es decir, una conexión de tipo normativo y no causal.

Cuando se habla de excepciones a la regla de exclusión probatoria, tradicionalmente se piensan en excepciones que involucran cuestiones causales. Así, se observa especialmente cuando se analiza el problema de la eficacia refleja. Tuvo su origen en la década del 20 del siglo pasado en los Estados Unidos y se cristalizó en *la regla del fruto del árbol envenenado*. Esto implica que si una prueba es obtenida de forma ilícita no sólo se impone la no admisión de esta prueba, sino de todas aquellas que se deriven de ésta. En España se ha reconocido que esta doctrina (que ya habían introducido de forma previa los tribunales) se recoge en el artículo 11.1. de la LOPJ, que dispone: «no surtirán efecto las pruebas obtenidas, directa o indirectamente, con violación de los deberes fundamentales» (Miranda, 2010).

Cuando Miranda (2010) analiza las excepciones, que señalaba hace más de una década en incremento³, podemos concluir que muchas de ellas refieren a conexiones o a un análisis en términos exclusivamente causales. Así, por ejemplo, se analiza como una excepción a la regla de exclusión la buena fe policial. Este concepto que parece indicar un concepto normativo, no lo es porque se refiere al error de hecho o de derecho: los policías pensaban que actuaban bajo una orden judicial legítima. Si bien podría conceptualizarse, en apariencia, en términos de conexión normativa, es estrictamente causal porque refiere a lo que los policías creían de hecho. La sentencia se sustenta en la conocida tesis del *deterrent effect* que persigue disuadir a los funcionarios policiales a que obtengan pruebas por medios ilícitos, declarándolas inadmisibles. Apunta a la ratio de la norma de exclusión probatoria, que implica concebirla como dirigida a las autoridades policiales o del Estado.

Dentro de las excepciones a la eficacia refleja, el mencionado autor señala la doctrina de la *fente independiente*, el *descubrimiento inevitable* y el *nexo causal atenuado*. Todas estas excepciones refieren exclusivamente a la conexión causal. La fuente independiente más que de una excepción involucra el propio alcance de la eficacia refleja,

³ Esta tendencia, como lo señala Muñoz (2023), ha ido en aumento. Recientemente un fallo del Tribunal Supremo Español considera que cuando los funcionarios policiales actúan sin que su actuación esté en el marco de un proceso penal (en un sentido muy amplio) que se centre en la figura de un acusado, es decir, no están buscando pruebas, tampoco rige la regla de exclusión probatoria. En el caso se trataba del extravío de una mochila con un soporte informático que uno de los policías encargados de recepcionar denuncias copió sin autorización judicial. Los archivos contenían imágenes de pornografía infantil.

porque si la prueba que es usada está en desconexión causal total con la prueba ilícita, el problema no se presenta de forma genuina. Que el autor lo mencione como desconexión causal, exime de cualquier otro comentario acerca de que el análisis se hace desde este punto de vista. El descubrimiento inevitable también hace referencia a una conexión causal o estrictamente de hecho. Neutraliza los efectos del contagio, cuando la nueva prueba que deriva causalmente de la prueba ilícita podría haber sido descubierta sin la mediación de la prueba inadmisibles. La del nexo causal atenuado es una variación de la de fuente independiente y se ha entendido de aplicación cuando, por ejemplo, se da una confesión días después con el sujeto en libertad y se le advierten previamente sus derechos. Miranda (2010) es crítico con estas excepciones porque considera que han sido utilizadas para justificar la no aplicación de la eficacia refleja de la regla de exclusión probatoria.

A las tradicionales excepciones que corren por el carril de la conexión de la causalidad, se podría pensar que la *conexión de antijuridicidad* va por el camino de encontrarle una explicación normativa a la conexión de las pruebas. Sin embargo, esto está lejos de ser así. La tesis del descubrimiento probable, que Miranda (2010) señala como parte de la doctrina de la conexión de antijuridicidad, da cuenta de que la teoría sigue poniendo énfasis en aspectos que son fácticos.

La teoría tuvo su origen en la Sentencia 81/1998 del Tribunal Constitucional (STC 81/1998, de 2 de abril de 1998) donde consideró un caso de narcotráfico en que se había intervenido un teléfono, pero, aparentemente, la decisión judicial de hacerlo si bien existió, no estuvo debidamente justificada porque sólo habían sospechas infundadas y no existía una justificación normativa para decretar la intervención del teléfono. Esta intervención es declarada ilegal en forma posterior. Sin embargo, el Tribunal rechaza aplicar el efecto reflejo de la prueba ilícita. Cuando se interfieren las comunicaciones por medio de las escuchas telefónicas, las autoridades policiales se enteran de una cita entre el acusado y otra persona. Se dirigen al domicilio donde se iba a dar el encuentro y por lo que surge de la sentencia, se da un enfrentamiento entre la policía y el acusado, donde este resulta herido y arroja un paquete con droga. El tribunal considera que tanto el hallazgo de la droga, como la declaración testimonial de los funcionarios policiales que participaron en el operativo son prueba válida que permite condenar al acusado.

El tribunal fundamenta la validación de estas pruebas en base a la teoría de la *conexión de antijuridicidad*, sosteniendo que, si bien las pruebas estaban conectadas causalmente, no lo estaban desde el punto de vista jurídico, por dicha razón se establece la conexión de antijuridicidad como fundamento de la teoría. Sin embargo, a poco que se analizan dichos fundamentos, no es que se descarte la conexión jurídica de las pruebas, sino que lo que se analiza es la violación del derecho fundamental en cuestión. De lo que se trata es de conceptualizar que en aquellos casos donde la violación del derecho fundamental no es grave, la prueba será admisible. Lo que considera el tribunal es que las llamadas se interceptaron con orden judicial, aunque esta orden no estuviese justificada desde el punto de vista jurídico. Se analiza también la

existencia o no de intencionalidad o culpa grave en la violación original del derecho fundamental. Pero parece claro que las pruebas analizadas, no sólo están conectadas desde el punto de vista causal, sino que también lo están desde el punto de vista jurídico. Porque quienes incumplieron con las normas jurídicas siempre son agentes estatales en el marco del proceso de recolección de pruebas.

En los casos que nos ocupan, la diferencia entre causalidad y normatividad pasa por lo siguiente: desde el punto de vista fáctico o causal, sin el hurto de Falciani o sin la copia presuntamente ilegítima de Manuel, ninguna de las Administraciones Tributarias hubiese tipificado la infracción o el delito de fraude fiscal. Sin que ocurriera el primer hecho (la copia informática) no se hubiesen producido los segundos hechos: las actuaciones de recolección de pruebas de las Administraciones tributarias. Sin embargo, si concebimos la prueba como atravesada por las reglas institucionales y las conceptualizamos como tales en el marco de los distintos procedimientos judiciales en el que se dan, la conexión causal no es o no implica de por sí una *conexión normativa*. Es imprescindible pensar en las normas que regulan la obtención de la prueba en el marco del procedimiento en el que están insertas.

En el caso tributario esto es especialmente importante porque las Administraciones tributarias suelen tener procedimientos administrativos regidos por reglas particulares que regulan cómo obtener las pruebas y porque están sometidas también a determinadas obligaciones que pueden incidir especialmente en el tema que nos ocupa. Es bastante común que los ordenamientos jurídicos impongan a los funcionarios públicos la obligación de investigar hechos en apariencia delictivos y de tratar de cierta forma a hechos que conocen por diversos medios.

De la sentencia analizada no surge cómo se enteró la AEAT del allanamiento de la morada en Francia de Falciani (Belarmino) por parte de la policía, no obstante, podemos suponer que tuvo conocimiento por la prensa. Dado este hecho, que llegó al conocimiento (aun cuando fuera informal) de la Administración española, recae sobre ella la obligación de disponer de las instrucciones necesarias para averiguar la verdad de los hechos y pesa sobre ella la carga de la prueba. Tiene la obligación de descubrir la verdad y para ello, como expuse al comienzo, tiene facultades especiales que son más potentes que las de cualquier particular. Tiene suscrito un convenio con Francia para evitar la doble imposición y el fraude tributario. Es en uso de esta facultad legal y legítima que solicita la información. Pero, además, todo indica que estaba obligada a hacerlo. Pesaba sobre ella determinada obligación. Cuando comienza el procedimiento probatorio, la Administración esperaba poder cumplir con la obtención de pruebas mediante cauces o canales diferentes que permitieran corroborar la hipótesis que originalmente la Administración se había planteado en cuanto a que Darío no declaró activos imponible para el impuesto a la renta de las personas físicas. Como reconoció el propio encargado de la unidad tributaria había otras hipótesis alternativas. Podemos suponer, por ejemplo, que los fondos no fuesen de Darío, sino sólo que estuviesen a su nombre pero que, en realidad, fueran de otra persona. Sin embargo, las sucesivas incomparecencias de Darío (fue citado once veces) impidieron la cons-

trucción de prueba diferente por parte de la AEAT. Ahora bien, la ilicitud de Falciani no guarda una relación jurídica con la AEAT, en el sentido de que no se trataba de un funcionario público actuando a su servicio, como reafirma el Tribunal Supremo.

En el caso uruguayo la situación es similar. Como ya se dijo, sin la denuncia de Manuel, la Administración Tributaria Uruguaya (DGI) no hubiese podido adentrarse ni conocer, probablemente, la maniobra defraudatoria del frigorífico. Pero también queda claro que no hay una conexión jurídica con el obrar de Manuel (exemplado) y la Administración, en el sentido de que no se trataba de un funcionario de la DGI, ni estaba actuando en colaboración con la Administración. La Administración Tributaria Uruguaya está obligada a recepcionar denuncias. No está dentro de sus facultades negarse a recibirlas. Una conducta de este tipo podría hacer a los funcionarios pasibles de falta administrativa grave o muy grave. Resta por verse si la recepción de la denuncia se hizo en legal forma, lo que analizaremos en la sección 6.2.

6.1. Las acciones y las descripciones de las acciones: el intento naturalista de Davidson que coadyuva a reafirmar la distinción entre causalidad y normatividad

La normatividad, a diferencia de la causalidad, no es una cuestión de causas, sino de razones. La relación de un hecho con otro hecho es de tipo causal. Dentro de los hechos, un tipo que al derecho le importa especialmente, son las acciones intencionales, es decir, la conducta de los sujetos. Davidson (1995) se inserta en una tradición analítica que no está preocupada especialmente por las cuestiones normativas sino más bien por los asuntos de la filosofía del lenguaje. Sus estudios de teoría de la acción están impregnados por esta tradición analítica en la que el autor se enmarca. Pese a plantear que las cuestiones de causas y razones no corren por carriles divergentes, entiendo que las valiosas distinciones que en materia de acción realiza Davidson, constituyen aportes relevantes que pueden ser utilizados en teoría de la prueba.

Davidson plantea que las acciones intencionales son un tipo especial de eventos y que para poder explicarlas necesitamos apelar a la noción de razones, que nos permita comprender qué fue lo que llevó al agente a hacer lo que hizo. Esto se debe a que presuponemos lo que el autor llama el Principio constitutivo de la racionalidad que sostiene que los agentes son racionales.

Es posible, sin embargo, apelar no ya a razones para entender una acción intencional sino a causas. Este tipo de explicación sostiene que para comprender determinado evento lo relevante es constatar cuál fue su causa. Una causa no justifica un evento porque no es una razón para que el evento suceda. En la relación causal no se constata ninguna relación lógica, sino que lo que se da es una regularidad en el precedente. Caorsi (2008) destaca que desde Hume es común oponer ambos tipos de explicaciones: la explicación basada en razones a la explicación basada en causas.

Pero la tesis de la interpretación radical de Davidson (1995)⁴ involucra que los deseos y las creencias del agente y su preferencia tengan una relación causal. Se verifica una relación causal entre eventos que son mentales (creencias, deseos, actitudes en general) y eventos físicos (conductas). Esta relación también se da a la inversa porque suele ocurrir que eventos físicos den origen a eventos mentales.

En el artículo «Acciones, razones y causas», Davidson (1995) batalla contra la tesis clásica de oponer racionalizaciones a explicaciones causales. Se propone defender la tesis que considera antigua y de sentido común acerca de que la racionalización es una especie de explicación causal. Las razones son también causas de las acciones. Una racionalización tiene dos componentes: a) justifica la acción al exponer la razón que tuvo el agente para realizar la conducta y b) explicita porqué la acción aconteció.

Davidson (1995) intenta confutar el argumento de que «una causa tiene que ser “lógicamente distinta de su supuesto efecto”; pero una razón de una acción no es lógicamente distinta de la acción; por tanto, las razones no son causas de las acciones», (p. 18) para lo cual descompone las razones de los agentes en dos partes. En primer lugar, una actitud favorable respecto de determinados tipos de acciones y, en segundo lugar, la creencia que su acción pertenece a ese tipo de acciones. Son actitudes favorables para Davidson los deseos estables o recurrentes del sujeto, pero también cualquier pasajero capricho. El autor pone el ejemplo de la acción intencional de encender la luz. Por un lado, tenemos que a) el sujeto tenía una actitud favorable de encender la luz (quería hacerlo), y a su vez b) tenía la creencia de que accionado el interruptor, la luz se encendería (se verifica la creencia de que accionar el interruptor es una clase de acción que pertenece a la clase de acciones «encender la luz»). La suma de la actitud favorable y la creencia, configuran lo que Davidson llama la *razón primaria* del agente.

Supongamos que en el ejemplo del interruptor y la luz, sin saberlo, al mismo tiempo, también alerto a un merodeador de que estoy en la casa. Esta acción es única, aunque haya dado cuatro descripciones de ella (presionar el interruptor, encender la luz, iluminar el cuarto, alertar al merodeador). El acto de presionar el interruptor se racionaliza por mi deseo de encender la luz, pero haber querido encender la luz no es una razón para alertar al merodeador. Las razones operan como racionalizaciones según las descripciones que se den de las acciones.

Esas cuatro descripciones de la misma acción, nos indican que una misma acción puede ser susceptible de más de una descripción. Que encendiera la luz es la causa de que alertara al merodeador, pero no es la razón por lo que lo hice. Me equivoqué si sostengo que la explicación de que: «Accioné el interruptor porque deseaba encender la luz», es equivalente a la explicación «Alerté al merodeador de mi presencia en la casa porque deseaba encender la luz», pese a que sean la misma acción. Esto tiene que ver con que una acción puede ser descrita de varias maneras y una razón no racionaliza

⁴ La interpretación radical es la interpretación de las preferencias de un hablante de una lengua que es totalmente desconocida para nosotros y que queremos entender y los desafíos asociados a dicha situación.

una acción sino más bien una cierta *descripción* de la acción. Precizando la noción de razón primaria, Davidson va a sostener que algo es una razón primaria por la que el agente realiza esa acción, bajo determinada descripción.

Pero la razón, para ser la razón primaria en los casos de acciones intencionales, que se distingue de los casos de justificaciones *ex post facto*, está en que dicha razón tiene que ser la razón por la cual la acción fue realizada. Davidson considera que las razones primarias de una acción son su causa.

El problema se presenta porque si consideramos que «encender la luz» y «alertar al merodeador de mi presencia en la casa» son la misma acción, y que al mismo tiempo lo primero es intencional, mientras que lo segundo no lo es, nos enfrentamos a la paradoja que la misma acción sería intencional y no intencional al mismo tiempo. El problema se despeja si consideramos que de lo que se trata es de descripciones diferentes de la misma acción. Según una de las descripciones el evento se presenta como racional y según la otra descripción no lo es. De lo anterior se deriva que lo que son intencionales no son las acciones, sino algunas de sus descripciones.

En su ensayo «De la acción», Davidson (1995) sostiene que actitudes y creencias son modos de describir ciertos eventos mentales, pero que bajo otras descripciones esos mismos eventos pueden ser descriptos como subsumidos en ciertas leyes. Pone el ejemplo de Hamlet:

Hamlet mata intencionalmente al hombre detrás del tapiz, pero no mata intencionalmente a Polonio. Sin embargo, Polonio es el hombre detrás del tapiz .. el matar de Hamlet al hombre detrás del tapiz es idéntico a su matar a Polonio (p. 66).

Hamlet realizó la acción intencional de matar al hombre detrás del tapiz, pero no la acción intencional de matar a Polonio. Si variamos el ejemplo y analizamos la situación donde por accidente Hamlet tropieza y mata a Polonio, en realidad esta no es una acción intencional de Hamlet, sino algo que simplemente le sucedió. Un mismo evento puede ser descrito en base a explicaciones causales según determinadas descripciones de esos eventos. De esta forma, la racionalización de una acción y su explicación causal corren juntas: son descripciones distintas de los mismos eventos. Si describimos ciertas acciones como eventos en términos de deseos y creencias, la acción se implica lógicamente de un modo determinado. Pero si los mismos eventos se describen físicamente explican de forma causal la acción bajo una descripción diferente (Caorsi, 2008).

En mi opinión, cuando se habla de deseos y creencias como actitudes proposicionales, la figura del agente involucra la cuestión normativa y la noción de razones. Las acciones se explican en términos de deseos, creencias y fines del agente, y estos conceptos no pueden ser introducidos sin apelar a la noción de normatividad. Explayarnos sobre esto excede el marco del presente artículo⁵, sin embargo, la contribución con el problema de la prueba se hace bastante claro.

⁵ Para ver con claridad lo que involucra la cuestión normativa puede consultarse a Korsgaard (2008).

La prueba, entendida como enunciado probatorio: *Falciani hurtó los archivos informáticos* puede ser *descripta* o concebida como ilícita o lícita según el contexto en el que es emitido el enunciado probatorio. A su vez, el enunciado: *Manuel copió los archivos de la doble contabilidad del frigorífico*, podría considerarse ilícito o lícito también según el contexto en el que se inserte, en términos de descripción de acciones. Vamos a profundizar en esto en el apartado siguiente.

6.2. La causalidad y la normatividad en las reglas de exclusión probatoria a la luz de los casos tributarios bajo estudio

La lista Falciani en tanto documento era el resultado de un ilícito, bajo una determinada descripción. Belarmino había obtenido la lista hurtándola de la base de datos del banco, por lo cual fue denunciado penalmente. En el marco del proceso de detención de Falciani era una de las pruebas clave que lo incriminaba. La lista, aunque ilícitamente obtenida, fue la prueba para incriminar a Falciani. Quiere decir que como prueba en el proceso penal contra Falciani desplegaba efectos. La ilicitud de la prueba sólo puede ser comprendida en el marco de un contexto particular y bajo una *descripción* particular. Como prueba del delito cometido por Falciani no es ilícita. Ahora bien, si Falciani pretendiera oponerla como prueba en contra de sus empleadores (el banco HSBC), por hipótesis, en el marco de un proceso de acción de reparación de daños, la prueba podría considerarse ilícita.

La lista Falciani es hallada en el domicilio en Francia donde se encontraba Falciani, en virtud de una requisitoria realizada por la policía suiza a la policía francesa. El allanamiento que realiza la policía de Francia es perfectamente legítimo y es realizado con arreglo a derecho. La prueba que obtiene la policía es lícita. Si bien la lista es única, se trata de un único *hecho*, podemos hablar de diferentes «pruebas» o «descripciones de pruebas», aunque el medio probatorio o fuente probatoria sea el mismo. Esto se debe a que la prueba, es la prueba de los hechos, no los hechos brutos y la prueba está atravesada por reglas constitutivas. El mismo hecho o fuente probatoria puede desplegar determinados efectos para un proceso (la condena a Falciani por hurtar la información), y a esos efectos ser prueba lícita, y puede, al mismo tiempo, eventualmente, ser una prueba ilícita si Falciani la quiere usar a su favor para un proceso de cobro de pesos a su empleador (el Banco HSBC).

Ahora bien, si Falciani utiliza la lista o hace referencia a ella como medio probatorio, también en determinadas hipótesis, podría hacerlo de una forma lícita o con apego a las reglas. Por ejemplo, si el HSBC alegara que hurtó muchos más datos que los que realmente hurtó, Falciani podría oponer la lista y decir «sólo hurté lo que figura en dicha lista y esa es la prueba de mi delito», y lo haría en ejercicio de su legítimo derecho constitucional de defensa.

Cuando observamos la conducta (la acción intencional de Falciani al hurtar los datos) parece claro que su acción es intencional con respecto a obtener información

de la que pudiera sacar un beneficio económico (la hipótesis manejada por las autoridades es que buscaba vender la información a otro banco). Pero también podría haber copiado la información a otros efectos, por ejemplo, obtener un respaldo de su participación en determinadas maniobras ilícitas. Su acción admite más de una descripción (siguiendo a Davidson) de acuerdo a su intención.

Un caso similar sucede con Manuel, el jefe de administración del frigorífico uruguayo. Manuel era un funcionario que se prestaba para ejecutar las maniobras defraudatorias encomendadas por sus empleadores y partía la facturación de la empresa, facturando sólo el 50 por 100 o el 25 por 100 de lo que vendía. Surge de la sentencia que fue condenado por un delito continuado de estafa. Engañaba económicamente a sus empleadores. Pero al ser partícipe de la maniobra ilícita de defraudación tributaria en calidad de tercero (para beneficiar a otro y que pagara menos impuestos de los que legalmente correspondía) Manuel se «ensuciaba» las manos con un ilícito. Si guardaba un respaldo de su maniobra delictiva podría tener algunas razones lícitas para hacerlo. Por ejemplo, delimitar hasta donde había llegado su participación delictiva y probarlo si era acusado por la Administración Tributaria (cuestión que efectivamente sucedió). Esto es, ejercer su derecho de defensa constitucional. Es más, efectivamente de la sentencia de la condena por estafa, surge que Manuel intentó defenderse de esta acusación en base a la prueba de la maniobra defraudatoria de la empresa, lo mismo sucede en los juicios laborales por notoria mala conducta. Del contexto de la sentencia puede inferirse que una estrategia posible para defenderse de la acusación de la estafa era sostener que en realidad él no estafó a la empresa, sino que lo que hubo fue una defraudación tributaria, de la cual él participó para beneficiar a la empresa que ahora lo denuncia. Pero, más modestamente, podemos quedarnos con la idea que, en tanto copartícipe de la maniobra defraudatoria, Manuel tiene derecho a guardar prueba de respaldo que deje en claro cómo fue su participación en el delito. Alguien que roba a una empresa, por ejemplo, haciéndose depósitos en su cuenta bancaria que no le corresponden parece que tiene derecho a probar, frente a una acusación, que hurtó 100.000 dólares y no 1.000.000.

Manuel realizó una denuncia a la DGI del Uruguay en su condición de particular y para ello adjuntó la prueba de que disponía. La mayoría del Tribunal reprocha a la Administración que la prueba no se recogió en legal forma y que esto la tornaría ilícita por su violación al artículo 12 de la Constitución uruguaya que consagra la regla del debido proceso, que abarca las garantías del derecho de defensa. En el caso, el Tribunal calificó de ilícita la prueba documental aportada al no preservarse la cadena de custodia. Sin embargo, incurre en una confusión en la lectura de los hechos.

Una alternativa podría haber sido que la Administración tributaria citara a los representantes legales del frigorífico en la instancia de denuncia, tal como pretende la mayoría del tribunal en su sentencia. Sin embargo, esto es muy discutible porque, si bien el derecho uruguayo no contiene normas que regulen cómo deben recepcionarse las denuncias, es pacíficamente admitido que se presenten de manera anónima para proteger al denunciante siempre que se garantice el derecho de defensa.

Sin embargo, el tribunal no advierte que tal comparecencia no le aporta ninguna garantía a la empresa en cuanto a la fidelidad de la fuente o procedencia de la documentación aportada en la denuncia. Lo único que se hubiese conseguido en esta instancia es confirmar que la prueba agregada por Manuel provenía de su computadora y fue incorporada al expediente. Pero Manuel podría haber sido un demente o un empleado mal intencionado que hubiese «plantado» toda esa documentación simulando que se trataba de una documentación de la empresa. Por lo que, en dicha instancia, no era posible asegurar ninguna cadena de custodia, ni la incolumidad de la prueba, en virtud de que nadie había presenciado el momento en el que Manuel copió dicha prueba desde la computadora u ordenador del frigorífico a su propia computadora.

Como otro problema conceptualmente independiente, el tribunal comete un error de hecho cuando le reprocha a la Administración que la documentación tendría que haber sido extraída de la computadora de la empresa y en presencia de sus representantes y haber sido individualizada. En sus propios resultandos (antecedentes de hecho) señala que el frigorífico entregó esta documentación a través de su representante y lo documentó en acta de fecha 9 de septiembre de 2011, como surge del Considerando III de la sentencia dictada en mayoría (STCA 170/22, de 17 de marzo de 2022). La documentación, como se trataba de las facturas de ventas oficiales de la empresa, no necesita estar individualizada. No son documentos informales, o documentos «en negro» (poseen número y su identidad está garantizada en los propios documentos). La visita a la empresa y la extracción de documentación por parte de la DGI se da precisamente para verificar la veracidad de la denuncia.

Lo que la sentencia en mayoría quiere decir o sugiere con la cita a la doctrina del *fruto del árbol envenenado* es que la prueba original es ilícita porque fue copiada por un ex empleado sin la autorización de la empresa. Más allá de que en el caso, por las razones antes expuestas, la prueba no podría considerarse ilícita (por el derecho de defensa que tiene Manuel, en tanto partícipe del ilícito), aun cuando así la concibiésemos, esta conclusión no puede trasladarse al actuar de la Administración tributaria que lo único que hace es recibir una denuncia y adjuntar la prueba que acompaña dicha denuncia, dentro del marco de las reglas jurídicas que regulan su actuar. Si se considera la prueba ilícita, la ilicitud no debería contagiar al actuar administrativo, en la medida en que la Administración actuó en el marco de sus competencias legales. Otorgó varias vistas (oportunidades de defensa) para que la empresa ofreciera contraprueba. También obtuvo prueba de terceros (una empleada y la empresa proveedora del software) de la maniobra defraudatoria del frigorífico y oportunamente citó a los representantes de la empresa para que analizaran la prueba aportada por el denunciante.

Si el archivo informático puede considerarse prueba ilícita, por hipótesis, en un reclamo laboral de cobro de pesos de Manuel contra el Frigorífico, no debería considerarse de la misma forma si se comporta como prueba que acompaña una denuncia al órgano recaudador de impuestos y este actúa en el marco de sus competencias lega-

les cuando la recibe. Tampoco podría serlo si oficia como la prueba de la comisión de su ilícito de estafa, o si es la prueba de su actuación como defraudador de impuestos, en calidad de tercero.

A los efectos que aquí importa no se trata de la misma «prueba» desde el punto de vista *normativo*, si existen reglas diferentes para su incorporación al procedimiento y si las finalidades que persiguen los sujetos que la incorporan a un procedimiento no es la de cometer ningún ilícito sino cumplir con sus obligaciones legales.

Darío, el contribuyente que omitió declarar sus cuentas bancarias, no tiene derecho a la intimidad de dichas cuentas con la Agencia tributaria española, y no lo tiene porque tiene la obligación de declararlo e incumplió con su obligación jurídica. Cuando la AEAT dispara el procedimiento de aplicación del convenio con Francia, actúa en el marco de sus competencias legales y la prueba agregada producto de estas actuaciones no debería considerarse ilícita porque su modo de obtención es con apego a las reglas. Quien, en todo caso, incumplió con las normas fue Belarmino, no la AEAT, como sostiene el Tribunal Supremo.

La conexión de causalidad corre por un carril y la conexión de normatividad por otro. Aunque ambas pruebas se conecten causalmente porque sin una no se hubiese producido la otra, otro razonamiento se impone en materia de *normatividad*. La ilicitud sólo se contagia si estamos en el marco del mismo proceso, bajo las mismas reglas.

7. CONCLUSIONES

1. Las reglas constitutivas o institucionales y la distinción entre acción y *descripción* de acción, nos ayudan a distinguir que una misma *fuerza probatoria* puede ser considerada como prueba ilícita o lícita, según los procedimientos en los que se inserta y las reglas que la regulan.

2. A los efectos de la eficacia refleja de la prueba ilícita, es necesario distinguir la causalidad de la normatividad, en el sentido de que las pruebas pueden estar conectadas causalmente pero no normativamente. Lo que la dogmática y la jurisprudencia llama conexión de antijuridicidad parece seguir poniendo el foco en cuestiones causales y de hecho y no normativas.

3. En el caso de la prueba tributaria, dado que las Administraciones Tributarias despliegan un procedimiento administrativo de obtención de pruebas contando con prerrogativas y competencias especiales, que suelen implicar facultades de fiscalización y de aplicación de convenios internacionales, se hace necesario y muy útil distinguir las cuestiones causales de las cuestiones normativas en el ámbito probatorio.

4. La distinción de estas dos esferas, la causal y la normativa, permite un contralor más racional de cuándo es necesario desechar una prueba por ilícita y cuándo no, y permite limitar la discrecionalidad del juzgador, impidiendo el establecimiento de sucesivas excepciones, sin que respondan a una línea argumental sólida.

BIBLIOGRAFÍA

- Aristóteles (1994). *Metafísica*. Gredos.
- Austin, J. L. (1991). *Cómo hacer cosas con palabras. Palabras y acciones*. Paidós.
- Caorsi, C. (2008). *Lenguaje, Acción e Interpretación. Una exposición crítica de la filosofía de Davidson*. Facultad de Humanidades y Ciencias de la Educación, Universidad de la República, CSIC.
- Davidson, D. (1995). *Ensayos sobre acciones y sucesos*. Crítica.
- Ferrer Beltrán, J. (2005). *Prueba y verdad en el derecho* (Segunda edición). Marcial Pons.
- Ferrer Beltrán, J. (2007). *La valoración racional de la prueba*. Marcial Pons.
- Gascón Abellán, M. (2010). *Los hechos en el derecho: bases argumentales de la prueba* (Tercera edición). Marcial Pons.
- González Lagier, D. (2010). *Quaestio facti. Nuevos ensayos sobre prueba y filosofía*. Palestra.
- Ho, H. L. (2021). The Legal Concept of Evidence. En E. N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (Winter 2021 Edition). <https://plato.stanford.edu/archives/win2021/entries/evidence-legal/>
- Holmes y Sunstein. (1999) *The cost of rights: why liberty depends on taxes*. WW Norton & Co.
- Hume, D. (2000). *Tratado de la naturaleza humana. Libro III. Acerca de la moral*. Eudeba.
- Korsgaard, C. (1996). *The Sources of Normativity*. New York: Cambridge University Press.
- Korsgaard, C. (2008). *The Constitution of Agency: Essays on Practical Reason and Moral Psychology*. Oxford University Press.
- Leiter, B. (2007). *Naturalizing Jurisprudence. Essays on American Legal Realism and Naturalism in legal Philosophy*. Oxford University Press.
- Mendonca, D. (2022). Ficciones legales. En L. Ramírez-Ludeña y J. Vilajosana (eds.) *Reglas constitutivas y derecho* (p. 129-139). Marcial Pons.
- Miranda Estrampes, M. (2010). La prueba ilícita: la regla de exclusión probatoria y sus excepciones. *Revista Catalana de Seguretat Pública*, 22, 131-151. <https://raco.cat/index.php/RCSP/article/view/194215>
- Miranda Estrampes, M. (2019). *Prueba ilícita y regla de exclusión en el sistema estadounidense: crónica de una muerte anunciada*. Marcial Pons.
- Moreso, J. y Vilajosana, J. (2004). *Introducción a la teoría del derecho*. Marcial Pons.
- Muñoz, A. (2023). La agonía de la regla de exclusión de la prueba ilícita. Variaciones sobre un caso trágico. Comentario de la sentencia del Tribunal Supremo, Sala 2ª, nº 891/2022, de 11 de noviembre. *Quaestio facti. Revista internacional sobre razonamiento probatorio*, 6, 301-352. https://doi.org/10.33115/udg_bib/qfi6.22909
- Osorio Morales, M. (2023). La prueba de los hechos en materia tributaria: una propuesta de desarrollo. *Dikaion*, 32(1), e3212. <https://doi.org/10.5294/dika.2023.32.1.2>
- Popper, K. (2003). *Conjeturas y refutaciones*. Paidós.
- Russell, B. (2013). *Lógica y conocimiento*. RBA.
- Searle, J. (1995). *La construcción de la realidad social*. Paidós.
- Searle, J. (2017). *Actos de habla: ensayo de filosofía del lenguaje*. Cátedra.
- Street, S. (2011). Mind-Independence Without the Mystery: Why Quasi-Realists Can't Have it Both Ways. En R. Shafer-Landau (ed.), *Oxford Studies in Metaethics* (Volume 6, p. 1-32). Oxford University Press.
- Tarski, A. (1969). *Verdad y demostración*. Publicaciones del Ciclo Básico Común, Facultad de Filosofía y Letras de la UBA.
- Taruffo, M. (2008). *La prueba*. Marcial Pons.
- Vázquez, C. (2013). *Estándares de prueba y prueba científica. Ensayos de epistemología jurídica*. Marcial Pons.

Legislación y jurisprudencia.

Constitución de la República (Uruguay). <https://parlamento.gub.uy/documentosyleyes/documentos/11/HTML>

España. Ley Orgánica 6/1985, de 1 de julio, del Poder Judicial. (BOE, núm. 157, 03/07/1985). <https://www.boe.es/eli/es/lo/1985/07/01/6/con>

Sentencia del Tribunal Constitucional (España) 114/1984, de 29 de noviembre de 1984 – ECLI:ES:TC:1984:114.

Sentencia del Tribunal Constitucional (España) 81/1998, de 2 de abril de 1998 - ECLI:ES:TC:1998:81.

Sentencia del Tribunal de lo Contencioso Administrativo (Uruguay) 170/22, de 17 de marzo de 2022.

Sentencia del Tribunal Supremo Federal (Alemania), de 14 de junio de 1960, BGHS 14, 358, 365.

Sentencia del Tribunal Supremo (España) 116/2017 (Sala de lo Penal), de 23 de febrero de 2017 (recurso 1281/2016) – ECLI:ES:TS:2017:471. [https://www.poderjudicial.es/search/contenidos.action?action=contentpdf&datasematch=TS&reference=7944729&links=por 10022116 por 1002F2017 por 10022&optimize=20170227&publicinterface=true](https://www.poderjudicial.es/search/contenidos.action?action=contentpdf&datasematch=TS&reference=7944729&links=por%2010022116%20por%201002F2017%20por%2010022&optimize=20170227&publicinterface=true)

CONJETURAS
Y REFUTACIONES

EPISTEMIC INJUSTICE IN THE CRIMINAL TRIAL: ENGAGING WITH GONZALES ROSE, HERDY, JALLOH AND OWUSU-BEMPAH

Federico Picinali*

ABSTRACT: Jasmine Gonzales Rose, Rachel Herdy, Tareeq Jalloh and Abenaa Owusu-Bempah have each written a paper commenting on my essay «Evidential Reasoning, Testimonial Injustice and the Fairness of the Criminal Trial», which appeared in *Quaestio Facti* in 2024. In this reply I engage with their insightful works. I discuss the advantages of framing in terms of «contributory injustice» the scenarios analysed in my original essay. I briefly study the conditions for the existence of a correlation between credibility excess and credibility deficit. And I provide the sketch of a theory of trial fairness, which I am currently developing elsewhere.

KEYWORDS: evidential reasoning; fairness; testimonial injustice; contributory injustice; credibility excess.

1. INTRODUCTION

I am very grateful to Jasmine Gonzales Rose, Rachel Herdy, Tareeq Jalloh, and Abenaa Owusu-Bempah for engaging with my essay «Evidential Reasoning, Testimonial Injustice and the Fairness of the Criminal Trial» (2024a). Their insightful and constructive responses build on the argument that I defended there, according to which testimonial injustice can render evidential reasoning (and hence, the criminal trial) unfair. This building work involved adding bricks, but also polishing and even

* I am very grateful to Tareeq Jalloh for comments on an earlier draft.

replacing some of the bricks I laid. The result is a construction that I had not fully envisaged at the time of writing my essay; a more solid and functional construction than the one I provided, to be sure. It is now my turn to engage with my readers' contributions with the aim to consolidate and enrich the construction further. Before doing so, however, a clarification is in order.

In her insightful paper, Rachel Herdy attributes to me the intention to provide a «general» theoretical framework; that is, a framework that could capture all significant types of testimonial injustice that may occur in the criminal trial. She correctly shows that my framework is not general in this sense, pointing out three types of testimonial injustice that the framework apparently does not capture. First, is «relational» testimonial injustice, where a credibility excess afforded to a party produces a credibility deficit for the other party, or where a credibility deficit suffered by a witness «ricochets» to the party that called the witness. Second, is agential testimonial injustice, where testimony is obtained from a speaker bypassing their epistemic agency, and is then given a credibility excess. Third, is testimonial injustice *simpli-citer*, where the testimony is not grounded in the stock of knowledge of the speaker qua member of a particular social group.

Herdy, however, seems to have misunderstood the goal of my essay. Only once in the essay I call my theoretical framework «general» (Picinali, 2024a, p. 204), and I am quick to qualify this attribute. My work builds on that of scholars—Gonzales Rose and Owusu-Bempah—who had already drawn a connection between testimonial injustice and racist evidential practices. I intended my framework to be «general» only in the sense that it would be applicable beyond the case of racism (indeed, the scenarios I discuss involve a wide range of identity prejudices). Still, given that the types of testimonial injustice mentioned in the previous paragraph are all acknowledged in my essay, one may ask why have I not attempted to offer a framework that could accommodate them. Why have I not, at least, discussed them in greater detail? As stated in the Introduction to the essay, my goal was to show that evidential reasoning—in particular, assessments of relevance and of probative value of individual items of evidence—is susceptible to an evaluation on grounds of fairness, not just to an evaluation on grounds of accuracy. To achieve this goal there is no need to provide a comprehensive taxonomy and analysis of the types of testimonial injustice that may occur in a criminal trial. It is sufficient to argue that *one such type* can occur in evidential reasoning, affecting speakers whose interests are crucial to trial fairness. I, therefore, focused on testimonial injustice that occurs when «the stock of knowledge that a party in the proceedings has qua member of a social group is ignored or discounted due to the adjudicator's identity prejudice against that group ... and, as a result, the party's argument about the relevance and the probative value of an item of evidence—argument that relies on such stock of knowledge—receives a credibility deficit» (Picinali, 2024a, p. 212). And I offered realistic scenarios where such an injustice takes place. Why have I picked this type of testimonial injustice, instead of other—perhaps less complex—types? As explained at 213, this type of testimonial injustice clearly evidences and reinforces a problematic

assumption that seems to characterise work in the rationalist tradition of evidence scholarship: the assumption that «there is only one stock of knowledge or, at least, that there is a stock of knowledge that is to be privileged for the purposes of adjudication» (Picinali, 2024a, p. 207). It is also due to this assumption that the rationalist tradition has failed to appreciate the moral dimension of assessments of relevance and of probative value. Namely, it has failed to appreciate that an oblivious and dogged adherence to the dominant stock of knowledge leads to ignoring or discounting other stocks; and that this, in turn, can render evidential reasoning unjust, rather than merely inaccurate. Such «ignoring or discounting» of stocks of knowledge are the crux of the type of testimonial injustice that I decided to highlight.

While my goal was, therefore, more modest than Herdy seems to have thought, her work—and that of Gonzales Rose, Jalloh and Owusu-Bempah—is important for the pursuit of the more ambitious goal of mapping, understanding and tackling the different types of testimonial—more generally, epistemic—injustice that may realistically occur in the criminal trial. The next two sections make contributions towards this goal, while the last section reflects on trial fairness.

2. TESTIMONIAL, HERMENEUTICAL ... CONTRIBUTORY

The type of testimonial injustice that I focus on in my essay has a marked hermeneutical component. According to Herdy, it originates from an underlying hermeneutical injustice, such that a «double» or «compound» (Herdy, 2024, p. 161-162) epistemic injustice apparently takes place. The credibility deficit is indeed caused by the partiality of the hermeneutical resources that the adjudicator employs to interpret the item of evidence. The adjudicator relies on the dominant stock of knowledge, ignoring or discounting the stock of knowledge that the party has qua member of a marginalised social group. The party suffers a credibility deficit precisely because their argument about relevance and probative value is based on a stock of knowledge which is not subsequently employed by the adjudicator in the assessment of such argument and, consequently, of the evidence.

While I acknowledge this hermeneutical component in the essay (fn 31), Herdy is right in suggesting that the component required more attention (Herdy, 2024, p. 159-163). Later I will come back to the question whether this component is best characterised as a hermeneutical injustice, but I have no reservation about Herdy's claim that the component is «at the root» (Herdy, 2024, p. 159) of the examples that I provide. Moreover, it is amongst the reasons why I singled out for analysis the type of testimonial injustice described in the previous section. If that type of testimonial injustice evidences the problematic assumption of the rationalist tradition discussed earlier, this is because of the hermeneutical imbalance that causes the credibility deficit. In fact, given my stated goal to show the possibility of injustice in evidential reasoning, one might go as far as to argue that it would have been sufficient for me to highlight the hermeneutical imbalance, instead of drawing attention to the

argument about relevance and probative value that a party may advance and to the credibility deficit that such argument may receive. This, I think, would be too fast. Considering only half of the picture—that is, focusing on the hermeneutical imbalance only—would have offered an impoverished understanding of the wrongs involved. In particular, it would have obfuscated the connection between the epistemic injustice and the unfairness of the trial. Let me clarify.

In the essay I argued that when a party is unable to participate effectively in the trial due to their testimony being unjustly discounted or pre-empted, the fairness of the trial is undermined. If fairness is indeed about participation—a point that I will come back to later—then a claim of unfairness requires a showing that participation is affected. While it is all but certain that the hermeneutical imbalance will hamper the participation of the party whose stock of knowledge is ignored—their testimony will be discounted or altogether pre-empted due to such imbalance—it is in the discounting and in the pre-empting of the testimony that the harm to participation is concretised. And such discounting and pre-empting are instances of testimonial injustice, not of hermeneutical injustice *per se*. So, I take Herdy's useful criticism as the argument that I should have said more about the hermeneutical aspect of my examples, not less about the testimonial injustice they involve.

My disproportionate attention to testimonial injustice is also problematic for another reason, addressed in Tareeq Jalloh's brilliant contribution (Jalloh, 2024). Jalloh draws attention to a challenge raised by Arcila-Valenzuela and Páez in a recent article that I referenced, but did not adequately engage with, in my essay (Arcila-Valenzuela and Páez, 2024). Arcila-Valenzuela and Páez argue that it is impossible to ascertain individual cases of testimonial injustice, due to the obstacles to proving the three essential elements of this form of epistemic injustice: an identity prejudice (understood as a stable personal trait), a credibility deficit and a causal link between the two. Let us assume that the challenge is sound. Even so, I don't think it detracts from the importance of studying testimonial injustice in the context of the criminal trial. By the authors' own admission, the existence of testimonial injustice is «undeniable» (Arcila-Valenzuela and Páez, 2024, p. 586), and it is evidenced by studies detecting «widespread credibility inequality» at the population level (Arcila-Valenzuela and Páez, 2024, p. 598): on average, members of some social group are believed less than members of some other group. In other words, if I understand them correctly, Arcila-Valenzuela and Páez claim that credibility inequality at the population level is sufficient evidence to conclude that individual instances of testimonial injustice do occur, even if no specific instance can ever be verified. Be that as it may, a credibility inequality at the population level is bad enough, even if it is the most that we can ascertain; and one should expect such inequality to affect also the criminal process, to the detriment of participants belonging to social groups that typically suffer from such inequality (racial minorities, women, those affected by disability etc.). If so, a discussion of the means

to prevent such inequality in the criminal process is surely valuable and topical, whether or not it is framed in terms of testimonial injustice.

Jalloh, however, offers another response to the «provability» challenge: to shift the focus from testimonial injustice to another type of epistemic injustice with different, and seemingly easier to prove, elements. This is contributory injustice, a form of epistemic injustice that I only touched on in the essay (fn 24). Contributory injustice occurs when there are different hermeneutical resources that the agent could use in an interpersonal epistemic enterprise—crucially, including resources that would allow a member of a marginalised group to conceptualise and to make visible to others their social experience—but the agent chooses to rely on structurally prejudiced hermeneutical resources, such that members of the marginalised group are thwarted in their contribution to the epistemic enterprise (Dotson, 2012, p. 31-32). In other words, the agent is wilfully ignorant of the hermeneutical resources of the marginalised group, and this effectively hinders the ability of members of such group to contribute to shared understanding. A significant difference between contributory injustice and hermeneutical injustice is that the latter—at least in its «central case» (Fricker, 2007, p. 147-152)—consists in a situation where marginalisation has produced a lacuna in the hermeneutical resources, such that the experience of the marginalised group is unintelligible to them (Dotson, 2012, p. 32; cf. Fricker, 2013, p. 1319). If so, contributory injustice is a better fit for the scenarios that I have discussed in my essay, where the epistemic agent at the receiving end of the injustice is able to make sense of their experience, while being unable to make it visible to the adjudicator. Moreover, unlike hermeneutical injustice, contributory injustice has a marked agential component. It is not merely a structural feature of the hermeneutical resources (say, the lacuna characterising hermeneutical injustice), but it is something that an agent (say, an adjudicator) does. This too makes contributory injustice a better fit than hermeneutical injustice for the cases I study in the essay, where the agency of the adjudicator plays a significant role. Finally, contributory injustice puts a stress on the hindrance that the wilful ignorance creates to the participation of marginalised agents in the epistemic enterprise. As pointed out earlier, this hindrance is where the unfairness of the trial materialises. A possible wrinkle with the suggested categorisation is that contributory injustice requires *wilful* ignorance (a choice to ignore), whereas the phenomenon I sought to describe encompasses cases of negligent overlooking of a party's stock of knowledge. Jalloh apparently finds no difficulty with subsuming cases of negligence under contributory injustice, as defined previously (Jalloh, 2024, p. 195). Should he be wrong in this regard, there would be good reason to extend the definition to encompass negligent ignorance.

Apart from descriptive accuracy, treating the phenomenon I study in the essay as an instance of contributory injustice, has the significant advantage of avoiding Arcila-Valenzuela's and Páez's challenge. Notice that contributory injustice—but the same goes for hermeneutical injustice—does not comprise an individual identity prejudice, a credibility deficit and a causal link between the two – these being the ele-

ments that are difficult to prove, according to the challenge. To be sure, both hermeneutical injustice and contributory injustice comprise an identity prejudice. But this is a «structural», rather than an individual, identity prejudice (Fricker, 2007, p. 155; Dotson, 2012, p. 29). It is not the stable personal trait targeted by Arcila-Valenzuela and Páez. It is, instead, a tilt in the hermeneutical resources, such that these «will tend to issue interpretations of [a] group's social experiences that are biased because insufficiently influenced by the subject group, and therefore unduly influenced by more hermeneutically powerful groups» (Fricker, 2007, p. 155). The proof of such tilt does not seem to raise formidable challenges.

Jalloh's shift of focus to contributory injustice is a clever move, and if contributory injustice is extended to encompass cases of negligent ignorance (rather than just of intentional ignorance) the proof of this injustice becomes even more achievable. The step from ascertaining the adjudicator's negligent ignorance of a stock of knowledge to concluding that a party suffered unfair treatment seems a short one indeed.

3. WHEN IS A CREDIBILITY DEFICIT THE CORRELATIVE OF A CREDIBILITY EXCESS?

Jasmine Gonzales Rose (2024) offers an insightful study of the testimonial injustice caused by credibility excesses – a topic also discussed by Herdy (2024) with particular focus on cases of «agential testimonial injustice» (Lackey, 2023). As stated earlier, this is a phenomenon that I acknowledged in the essay, but did not study in any detail. Gonzales Rose's analysis is rich and compelling, and it convincingly shows that an adequate understanding of the phenomenon of testimonial injustice in the criminal trial cannot prescind from a study of the effects of credibility excesses afforded to privileged speakers.

Gonzales Rose invites us to consider the familiar scenario in which a party in the process (say, the defendant) is from a racial minority and the other party (say, the prosecution) and the witnesses it calls (say, police officers) are White. Building on the work of José Medina (2011), among others, she argues that the credibility excess that the White party and witnesses are likely to be afforded automatically translates into a credibility deficit suffered by the opposite party. She calls this phenomenon a «transferred epistemic harm», to capture the indirect harm that occurs when a privileged speaker's credibility excess correspondingly reduces a subordinated speaker's epistemic ability or agency in a way that causes epistemic injury *qua* knower» (Gonzales Rose, 2024, p. 183). Gonzales Rose also clarifies that this phenomenon, while not unique to the criminal trial, is prevalent there due to the peculiar «credibility economy» (p. 175) of the trial, which she characterises as a «zero-sum game» (p. 180): to increase credence in a party's account means to decrease credence in the account of the opposite party.

Here I would like to clarify the features of criminal fact finding that enable the phenomenon aptly described by Gonzales Rose; in other words, the conditions due to which a credibility excess correlates with a credibility deficit. I start from a passage by Medina, which is used by Gonzales Rose as the springboard for her analysis.

Credibility is not assessed in the abstract ... but rather, in a comparative and contrastive way... So, those who have an undeserved ... credibility excess are judged comparatively more worthy of epistemic trust than other subjects, all things being equal; and this is unfair, not only to them but also to others who do not receive this privileged treatment, not because of a failure in equal distribution but because of a failure in proportionality, for the degrees of credibility given to subjects have to be proportional to their epistemic merits and the presumptions that apply to subjects in their situation. A credibility excess constitutes an epistemic injustice when and because it involves the undeserved treatment of an epistemic subject who receives comparatively more trust than other subjects would under the same conditions. The credibility excess assigned to some can be correlated to the credibility deficits assigned to others not because credibility is a scarce good ... , but because credibility is a comparative and contrastive quality, and an excessive attribution of it involves the privileged epistemic treatment of some ... and the underprivileged epistemic treatment of others (Medina, 2011, p. 20).

It seems that here Medina is making—and perhaps conflating—two distinct arguments. First, he argues that a credibility excess is an epistemic injustice for the mere reason that it causes disproportionality, across epistemic agents, between the respective epistemic merits and the levels of credibility they respectively receive. Notice that for this disproportionality to occur no one needs to suffer a credibility deficit: it is sufficient that, while most agents are assigned a credibility level that is reasonable given the available evidence and the agents' epistemic merits, some agents are assigned a credibility excess. There is a second argument in Medina's passage, and it is the argument that interests me here. Medina argues that a credibility excess *can* correlate with a credibility deficit (although in the following page he shows sympathy for the view that the correlation is not just possible, but very frequent, if not necessary). This means that believing an epistemic agent more than one should, may well involve believing another epistemic agent less than one should. This, according to Gonzales Rose (2024), is precisely what tends to happen in a criminal trial when a defendant from a racial minority is confronted by a White prosecutor and White prosecution witnesses.

While I don't think that a credibility excess always correlates with a credibility deficit, I am persuaded that the correlation generally exists in the context of the interaction between the parties of a criminal trial. And I suggest that this is due to the structure of the decision-problem with which a trial is concerned. Gonzales Rose writes of the credibility economy of the trial that it is a «zero-sum game», and she seems to consider this the reason for the occurrence of the correlation. Her account seems correct, but it falls short of clarifying why the credibility economy of the trial is such a game. A plausible clarification appeals to the consideration that, at the most general level of the trial contest, the parties defend hypotheses that are *mutually exclusive* and *jointly exhaustive*. These are the hypothesis of innocence and the hypothesis of guilt. They are *mutually exclusive* because they cannot both be true. They are *jointly exhaustive* because they fill up the universe of possibilities artificially created by the criminal law. The defendant is

either innocent or they are guilty: *tertium non datur*. In fact, someone who is open to a probabilistic understanding of criminal fact finding may well say that this decision problem is a «one-sum game», in that the sum of the probability of guilt and of the probability of innocence exhausts the entire probability spectrum, thus being equal to one.

Take away either of the two features mentioned, and the correlation between a credibility excess and a credibility deficit ceases to exist. To see this consider a couple of examples. In the first example, we have a historian who is seeking to answer the question of how a particular war started. The historian hears the testimony of a survivor claiming that the war was single-handedly and wantonly initiated by state X; but they also hear the testimony of another survivor claiming that it was single-handedly and wantonly initiated by state Y. The two hypotheses are mutually exclusive. But they are not exhaustive. There are other possibilities countenanced by the historian: perhaps the war was single-handedly and wantonly initiated by a third state; perhaps it originated from an accident, rather than deliberate action, etc. The historian gives a credibility excess to the survivor accusing state X, assigning to that hypothesis a probability of .6 when it only deserves a probability of .5. The historian can do so without also decreasing the credibility assigned to the survivor accusing state Y and the corresponding probability. This probability may remain constant at, say, .2. The credibility excess may simply involve a decrease in the probability of the (disjunction of the) other possible hypotheses (i.e., the responsibility of a third state, the accident, etc.). In the second example, instead, we have a scientific journalist who is writing a piece on what stars are made of. They interview a bunch of astrophysicists. A group of astrophysicists somewhat unhelpfully says that stars are made of some material, without being able to pin down any material. Another group says that stars are made of carbon dioxide. These two hypotheses are (indeed, the first alone is) exhaustive (at least if we exclude the option that stars are an optical illusion), but they are not mutually exclusive. So, the journalist may well assign a credibility excess to the group of scientists claiming that stars are made of carbon dioxide, without also inflicting a credibility deficit on the group of scientists who merely acknowledged that they are made of something.

The mutual exclusivity and the joint exhaustiveness of the hypotheses of innocence and of guilt mean that a credibility excess assigned to a party will translate into a credibility deficit assigned to the other. By the same token, the reverse is also true. As we are reminded by Gonzales Rose, US and UK societies and institutions are such that in a trial contest between a White party and a party from a racial minority, the White party will likely be afforded a credibility excess. This will, therefore, translate into a credibility deficit for the other party. If someone's concern is limited to the question of trial fairness, this analysis of the dynamics between credibility excess and credibility deficit within the trial may suffice. But Gonzales Rose also hints at the externalities of credibility excesses: the effects that excesses afforded in the trial may have beyond the trial itself. She writes:

When a White speaker's credibility is routinely boosted, this does not exist in isolation or as a one-time occurrence. Rather this is part of an epistemic-privileging routine and pervasive subordination which consistently and systematically exalts members of the White racial group to entrench their power and privilege to the corresponding and collective epistemic detriment of Black and other non-White people (Gonzales Rose, 2024, p. 179).

The credibility excess afforded to White speakers in a trial contest with non-White speakers does not only generate a credibility deficit for the latter and, hence, trial unfairness. It also contributes to entrenching and bolstering racist societal structures and dynamics, with both practical and epistemic ramifications. This recognition is important because it shows that even when a decision-problem is not characterised by mutually exclusive and jointly exhaustive hypotheses, a credibility excess afforded therein to a member of a privileged group may have an unjust impact outside that decisional context, possibly contributing to the infliction, in other epistemic enterprises, of credibility deficits to the detriment of the disadvantaged and oppressed.

4. FAIRNESS, PARTICIPATION AND EPISTEMIC INJUSTICE

In her rich contribution, Abenaa Owusu-Bempah (2024) criticises the English and Welsh law on the evidential significance of the defendant's silence, discusses possible remedies to testimonial injustice in the trial context and highlights forms of such injustice that I only touched on in my essay. These include the case, also discussed by Herdy (2024), of a credibility deficit suffered by a witness and «ricocheted» to the party that called the witness; but also, the case of testimonial injustice suffered by defence counsel, and the case of testimonial injustice suffered by the defendant at the hands of their counsel (Owusu-Bempah, 2024, p. 145-146). The emerging picture is that of a depressingly varied and realistic range of possibilities for epistemic injustice. Owusu-Bempah also raises some probing questions about the conception of trial fairness that I offered in my essay. Here I try to provide some initial answers to these questions, based on work that I am currently doing to develop that conception (Picinali 2024b).

Owusu-Bempah's critique focuses on the complainant and their participation. To begin with, she questions whether the complainant is properly referred to as a «party» and she suggests calling them a «participant» instead (Owusu-Bempah, 2024, p. 144). While in the essay I do call the complainant a «party», here I will use the term «stakeholder», so as not to appear to take for granted the complainant's entitlement to participate or their actual participation. In my essay I wrote that

for the purposes of my argument it matters not whether the complainant, like the defendant, has a right to a fair trial—that is, an actionable prerogative to a fair contest—or mere protected interests that are relevant to trial fairness—that is, prerogatives that are not actionable by the complainant, but that nonetheless should be respected, through providing appropriate treatment to the complainant, for the trial to be fair. Either way, under a pluralistic conception of fairness a trial cannot be fair if the complainant is not given the treatment that they are owed, given the legitimate interests they have at stake. Indeed, in these circumstances whoever has a right to a fair trial will suffer a breach of said right: they will be denied that to which they have a right. So,

even if the complainant had no right to a fair trial, the fact that they are not given the treatment they are owed—in particular, the opportunity for meaningful participation—means that the trial was unfair, and that the defendant suffered a breach of their right to trial fairness. While possibly counterintuitive for some readers, this strikes me as the inevitable implication of a pluralistic conception of fairness (Picinali, 2024a, p. 223-224).

Owusu-Bempah is unpersuaded by this passage, as the following excerpt from her article indicates. She writes:

Picinali states that it matters not for his argument whether the complainant has a right to a fair trial, or mere protected interests that are relevant to trial fairness ... But, if the complainant is not a party to the case and does not hold participatory rights, more needs to be said about: the nature and source of the participation owed to complainants; how testimonial injustice to the complainant (independent of the prosecution party) renders the trial unfair; and how testimonial injustice to complainants causes the defendant to suffer a breach of their right to trial fairness, a claim which Picinali recognises will seem counterintuitive to some (Owusu-Bempah, 2024, p. 144).

Here is a stab at answering these important questions. The institution of the criminal trial is based on values. These values may vary from one jurisdiction to another and from one historical period to another. Modern examples are the value of truth and the value of respect for the subjects involved. There can also be «second-order» values. The institution of the trial will likely involve a framework of rules designed to pursue some foundational values such as truth: respect for these rules is a second-order value on which the trial is grounded. Based on these uncontroversial ideas, I define a stakeholder as anyone who has a protected interest concerning the outcome and/or the conduct of the trial, where an interest is «protected» if its pursuit is supported by the values that inform the trial. So, an interest in a true outcome is protected if truth is indeed amongst such values.

Consider now the complainant. Typically, they have an interest in the conviction of the defendant. Is this a protected interest? Is the complainant, in other words, a stakeholder? The answers to these questions depend on which values inform the trial. If truth is one of them, the answers seem affirmative indeed. How so? It would seem that, unless we know that the defendant is guilty, we cannot conclude that the value of truth supports pursuing the complainant's interest in conviction. But if we knew this, the trial would lose much of its *raison d'être*, such that we would not have a complainant to speak of. In fact, here we must suspend judgement, lest we preempt the judgment of the adjudicator. At most we can say that the defendant *may be* guilty and, hence, that the complainant *may have* a protected interest in conviction. But this reasonable possibility suffices, in my view, to treat the complainant as a stakeholder. Similar considerations lead to the conclusion that the defendant too is a stakeholder. Of course, there likely are other stakeholders (e.g., the prosecution, the public). Moreover, complainant and defendant may have protected interests other than interests in outcomes (e.g., an interest in the respect of procedural rules). There is no need to explore these questions for present purposes.

The next step is to recognise that every stakeholder is entitled to participate in the trial in order to further their protected interests. For the criminal justice system to

stay true to the values that inform the trial, it must allow for the stakeholders' participation: participation is a significant (though, perhaps, not the only) means through which those interests and, hence, those values are furthered in any given trial. The modes of participation may vary from mere observation to examining witnesses or addressing the adjudicator. Moreover, not all stakeholders need to be afforded the same modes and extent of participation. Trial fairness, though, requires that for each stakeholder the same proportion holds between the weight of protected interests and the opportunities to participate. It follows that those with weightier interests should be afforded more significant opportunities to participate. How an interest is weighed and how proportionality between interests and participation is realised in practice are thorny questions that I cannot address on this occasion. The important message here is that anything that disrupts said proportionality causes unfairness. A testimonial injustice suffered by the complainant or by the defendant surely has that potential. If we assume that the law by and large realises, across all stakeholders, a proportionality between the weight of protected interests and opportunities to participate, then a testimonial injustice suffered by any stakeholder is likely to disrupt the proportionality and cause unfairness. A right to a fair trial is understood here as a right to that proportionality; it is therefore breached whenever and however the proportionality is disrupted. Of course, the law governing the current English and Welsh trial may not realise that proportionality, such that the trial may be unfair to begin with, irrespective of any testimonial injustice that were to take place...

This is not a theory, but the sketch of a theory. While it may not persuade Owusu-Bempah, I believe it gives some initial plausible answers to her questions.

BIBLIOGRAPHY

- Arcila-Valenzuela, M. and Páez, A. (2024). Testimonial Injustice: The Facts of the Matter. *Review of Philosophy and Psychology*, 15, 585-602. <https://doi.org/10.1007/s13164-022-00662-w>
- Dotson, K. (2012). A Cautionary Tale: On Limiting Epistemic Oppression. *Frontiers: A Journal of Women Studies*, 33, 24-47. <https://doi.org/10.5250/fronjwomestud.33.1.0024>
- Fricker, M. (2007). *Epistemic Injustice. Power and the Ethics of Knowing*. Oxford University Press.
- Fricker, M. (2013). Epistemic Injustice as a Condition of Political Freedom? *Synthese*, 190, 1317-1332. <https://doi.org/10.1007/s11229-012-0227-3>
- Gonzales Rose, J. B. (2024). A Critical Perspective on Testimonial Injustice: Interrogating Witnesses' Credibility Excess Trials. A Comment on Federico Picinali's «Evidential Reasoning, Testimonial Injustice, and the Fairness of the Criminal Trial» *Quaestio Facti*, 7, 173-184. https://doi.org/10.33115/udg_bib/qf.i7.23043
- Herdy, R. (2024). Testimonial Injustice in Evidential Reasoning: A Reply to Federico Picinali. *Quaestio Facti*, 7, 153-172. https://doi.org/10.33115/udg_bib/qf.i7.23031
- Jalloh, T. (2024). Understanding Epistemic Injustice as Contributory Injustice: A Comment on Picinali's Argument. *Quaestio Facti* 7, 187-199. https://doi.org/10.33115/udg_bib/qf.i7.23021
- Lackey, J. (2023). *Criminal Testimonial Injustice*. Oxford University Press.
- Medina, J. (2011). The Relevance of Credibility Excess in a Proportional View of Epistemic Injustice: Differential Epistemic Authority and the Social Imaginary. *Social Epistemology*, 25(1), 15-35. <https://doi.org/10.1080/02691728.2010.534568>

- Owusu-Bempah, A. (2024). Understanding and Preventing Testimonial Injustice in Criminal Proceedings: A Comment on Federico Picinali «Evidential Reasoning, Testimonial Injustice and the Fairness of the Criminal Trial». *Quaestio Facti*, 7, 139-152. https://doi.org/10.33115/udg_bib/qf.i7.23007
- Picinali, F. (2024a). Evidential Reasoning, Testimonial Injustice and the Fairness of the Criminal Trial. *Quaestio Facti*, 6, 201-235. https://doi.org/10.33115/udg_bib/qf.i6.22888
- Picinali, F. (2024b). What is a Fair Criminal Trial? *Unpublished Manuscript*.

TAKING THE EVOLUTION OF THE STANDARDS OF PROOF FOR A CRIMINAL CONVICTION SERIOUSLY*

Jacopo Della Torre

Associate professor of Criminal procedure

University of Genoa, Italy

jacopo.dellatorre@unige.it

ABSTRACT: The article offers a diachronic and comparative analysis of different standards of proof for a criminal conviction. The first part focuses on the attempt of medieval and early modern Roman-canon systems to clarify this type of rule through a network of legal proofs. The second part analyses the origins of the main standards for a criminal conviction used today: moral certainty, beyond reasonable doubt and *intime conviction*. The final part looks to the future, asking whether traditional decision-making criteria should be maintained or replaced by new ones based on the principles of contemporary epistemology.

KEYWORDS: standard of proof; reasonable doubt; moral certainty; *intime conviction*; legal proof.

SUMMARY: 1. INTRODUCTION.— 2. CONCEPTUAL AND TERMINOLOGICAL PREMISES.— 3. THE (SHATTERED) DREAM OF PRE-MODERN ROMAN-CANON SYSTEMS: THE RATIONALISATION OF STANDARDS OF PROOF THROUGH LEGAL PROOF RULES: 3.1. The standard of *luce meridiana clariores* and its functional connection with the legal

* I presented an earlier version of this work in Milan (Bocconi University, 7 March 2024). I also presented some of the ideas in this article at an online conference organised by the University of Nottingham on 28 November 2024. I am grateful to all the participants at these events who commented on my work. In particular, I would like to thank Diego Dei Vecchi, Paul Roberts and Giovanni Tuzet. I am also grateful to Paul Roberts for his hospitality at the University of Nottingham, which was essential for the English and US parts of the research. A previous version of this article appeared as Della Torre 2023a. Finally, I thank the reviewers of this journal for their constructive comments on the submitted draft.

proof rules; 3.2. The operational difficulties of the Roman-canon system of legal proof: the rise of the confession and the problem of the *indicia indubitata*.— 4. THE ORIGIN OF MORAL CERTAINTY AND REASONABLE DOUBT: 4.1. The theological background; 4.2. The Anglo-Saxon experience; 4.3. The Roman-canon experience.— 5. THE COLLAPSE OF THE SYSTEM OF LEGAL PROOF AND THE EMERGENCE OF THE *INTIME CONVICTION*: 5.1. The Enlightenment criticism of legal proof rules; 5.2. The rise of the two-faced rule of *intime conviction*; 6. THE AMERICAN EXPERIENCE AND THE ENGLISH DEVIATION.— 7. THE BARD'S AGE.— 8. CONCLUSIONS: WHAT STANDARD OF PROOF FOR A CRIMINAL CONVICTION FOR THE FUTURE OF CRIMINAL JUSTICE?— BIBLIOGRAPHY.

1. INTRODUCTION

In recent years, the issue of standards of proof has received increasing attention in both common law and civil law jurisdictions¹. As we shall see (§ 7), the reasons for this transversal interest are manifold. In some countries, the phenomenon has been encouraged by legislative attempts to explicitly codify evidentiary thresholds that were previously only implicit. In other cases, it has been the case law that has sought to better define the rules of decision-making. Finally, certain supranational institutions—such as the European Court of Human Rights—have also played an important role.

However, despite the wide range of the debate that has developed on this subject, it should be noted that the results obtained are rather heterogeneous and not without ambiguities. Suffice it to say that there is no unanimity even on some fundamental aspects, such as the definition of the concept of “standard of proof”², or on the real capacity of this type of rule to contribute to the proper functioning of legal systems. While there are those who are in favour of the primary importance of this category (Ferrer Beltrán and Tuzet, 2018, p. 455 ff.), there are also those who are of the opinion that standards of proof are nothing more than a relic of the past that we could well do without (Nieva Fenoll, 2020, p. 119 ff.). As we shall later see (§ 2), at the heart of these opposing visions lies a bitter dispute between those who, for legal and moral reasons, consider it possible and/or desirable to establish objective or intersubjective thresholds of proof, and those who, for various reasons, oppose this thesis.

The aim of this article is to take a step forward through a diachronic and comparative analysis dedicated to a particularly problematic category of standards: those aimed at determining the *quantum* of proof required for a criminal conviction. To understand the delicate nature of this issue, it is enough to recall that one of the most

¹ See Allen (2013, p. 43 ff.); Ambos (2023, p. 167 ff.); Clermont (2013); Ferrer Beltrán (2021b); Id. (2007, p. 149 ff.); Laudan (2006); Poli (2023); Nance (2016); Nieva Fenoll (2010, 85 ff.); Stella (2003, p. 116 ff.); Summers (2023, p. 264 ff.); Taruffo (2009, p. 218 ff.); Tuzet (2023, pp. 226-247 and 264-274).

² As Roberts and Zuckerman (2022, p. 267) point out «what, exactly, is a criminal standard of proof? This question is more complex, and its answer more controversial, than widely appreciated».

authoritative European scholars admitted that «*la degré de certitude qui conditionne la décision de culpabilité est une des questions les plus obscures du droit pénal*» (Delmas-Marty, 1996, p. 59).

The structure of the essay is as follows: after some introductory clarifications on the concept of “standard of proof”, which are intended to show how these rules fulfil several essential roles for the proper functioning of legal systems (§ 2), the first part focuses on its use in pre-modern and early modern Roman-Canon legal systems (§ 3). This part of the analysis has a specific purpose: to draw attention to one of the most ambitious attempts to make standards of proof less subject to the arbitrary power of adjudicators, perpetuated by the creation of a complex network of legal proof rules³. This was an experience that proved unsuccessful in the long run, but not for that reason unable of providing valuable lessons for the present day.

The second part of the article (§§ 4-7) contains a diachronic analysis of the origins and development of the principal standards of proof still used today for the imposition of punishment: moral certainty, beyond any reasonable doubt (hereafter BARD) and, finally, *intime conviction*. This will show that not only do these rules have very distant origins in time, but that they are also closely connected. And it is also because these rules are the fruit of cultural and legal changes that took place centuries ago that, as we shall see, interpreters today struggle so much to identify a minimal meaning of them.

Awareness of this fundamental historical fact will be the starting point for the last part of the work (§8), which will be devoted to understanding whether it is appropriate in contemporary legal systems to continue to maintain the traditional rules of decision-making, or whether, as some scholars argue, they should be replaced by new ones based on the principles of contemporary epistemology.

2. CONCEPTUAL AND TERMINOLOGICAL PREMISES

It is well known that the factual assertions made by the parties in legal proceedings are mere hypotheses which, as such, may be true or false (Taruffo, 2009, p. 218). Reducing this background uncertainty is one of the most delicate tasks of the law of evidence. By supporting «the factual claims made by the parties and the findings of fact made by the decision-makers» (Tuzet, 2021, p. 90), evidence ensures that the truth remains the point of reference that guides the direction in which the judicial process must move⁴. But when can a hypothesis about a fact be said to have reached such a level of corroboration that it can be considered sufficiently proven? To provide

³ For a recent English-language overview of the Roman-canon system of legal proof, see Damaška (2019).

⁴ On this point, see Caprioli (2017, p. 317 ff.); Ferrer Beltrán (2004); Summers (2023, p. 249 ff.); Taruffo (2009b); Ubertis (2021, p. 2).

a consistent and predictable answer to this question, legal systems have developed a specific set of rules: the standards of proof. To appreciate the long tradition of this type of rule, it is enough to remember that the medieval Roman-canon systems already used the category “*gradus probationis*”, which bears strong similarities to the modern concept of “standard of proof”⁵.

Having clarified this, it is possible to give an initial definition of the concept, which is necessarily broad because it aims to reconcile heterogeneous opinions: “standards of proof” are rules, either explicit or implicit, that specify the minimum threshold that must be reached for a hypothesis to be accepted as sufficiently proven by the trier of fact. Despite the vagueness of the proposed definition, it is useful for understanding the main functions of this category of rules in the legal context.

The first is to oblige public authorities to verify that a hypothesis has a certain degree of corroboration before it can justify the adoption of a measure that may affect a person’s legal sphere (Ferrer Beltrán, 2021b, 112 ff.). It follows that such rules act as a shield against the arbitrary restriction of individual rights by the authorities, protecting the most fundamental rights of the individual, such as the presumption of innocence, individual liberty, the right to property and, in some legal systems, even the right to life. Indeed, where there is a standard, the factfinder must operate within a “constrained discretion”, having to consider whether the prescribed threshold has been met⁶.

Secondly, standards also serve as an essential means of justifying the choices made by the decision-makers. Thanks to these legal rules, the fact-finders should be able to choose logically *ex ante* and, if necessary, to justify *ex post* why, among the hypotheses supported by the compendium of evidence, one should be accepted rather than another. It is therefore not by chance that many scholars have stressed the privileged relationship that exists between the standards of proof and the motivation of judgments (Damaška, 2019, p. 143; Ferrer Beltrán, 2021b, p. 109 ff.).

Thirdly, the decision-making rules act as an incentive for the parties, whether public or private, to seek and provide sufficient evidence to meet the individual thresholds set at the various stages of the procedure. This “driving force” of standards is particularly important for investigating authorities: these rules can indeed favour the conduct of complete investigations, both in terms of quantity and quality. And this clearly shows, from a privileged perspective, that there is a close link between the efficiency of justice and standards of proof (Della Torre, 2023b, p. 327 ff.). However, it should not be overlooked that these provisions also play an important role from a perspective more closely linked to the defence of the accused: indeed, knowing how much evidence the prosecution will have to produce in order to obtain a decision in its favour is essential information for making a rational decision on whether to

⁵ We find the Latin expression *gradus probationis*, for example, in Bartolo da Sassoferrato (1562, Comment on Book XII of the Digest., tit. II, *lex XXXI*, f. 564, § 20 et ff.).

⁶ On this point, see Iacoviello (2023, p. 300 f.).

articulate the defence strategy passively or actively, for example by agreeing to enter into a plea bargain (Ferrer Beltrán, 2021b, p. 113).

What has been observed illustrates how standards of proof are related in two ways to the rules that allocate the so-called risk of failure to prove the alleged facts between the parties⁷. On the one hand, they allow to know when the burden of proof can be considered to have been met. This makes it easy to understand why Anglo-Saxon systems use the concept of “burden of persuasion” in addition to the concept of “burden of production”: the former refers to the burden of persuading the trier of fact to the extent required by a standard of proof. At the same time, standards provide the material conditions for the application of so-called “decision rules” in the strict sense, i.e. those rules which resolve procedural uncertainty—i.e. the absence of the *quantum* of evidence required by the standard—in a manner favourable to one party or the other⁸.

However, it should be made clear that nothing requires the existence of a single standard of proof within a given jurisdiction or even within the same judicial procedure⁹. In this respect, it should be noted that standards of proof are applicable not only when the final judicial decision on the disputed facts is to be taken, but at all stages of the proceedings when it is necessary to verify, even provisionally, whether a minimum level of support for a given fact has been reached. From this point of view, standards can be divided into at least three macro-categories (Della Torre, 2023b, p. 333). There are: a) “propulsive standards”, which set out the evidentiary requirements that must be met in order to decide whether to prosecute or to proceed from pre-trial to trial¹⁰; b) “incidental standards”, which are used to decide whether or not to accept a case as proven at a particular sub-stage of the procedure (e.g. at the pre-trial detention stage¹¹); and, finally, c) “decision-making standards in the strict sense”, which are intended to guide the trier of fact in assessing whether there is sufficient evidence to support a conviction¹².

⁷ On the relationship between burdens and standards of proof, see Allen (2014, p. 195 ff.) and Stein (2005, p. 118 ff.).

⁸ The link between these categories of norms is that standards of proof determine the evidential requirements for the application of “decision rules”. And it is precisely because of this inextricable link that the terms “standard of proof” and “decision rule” are often used synonymously, even though they focus to a greater or lesser extent on different regulatory profiles.

⁹ See Ferrer Beltrán (2021b, p. 100 f.) who rightly insists that standards of proof should be built on an ascending scale, becoming more and more demanding. For a contrary view see Ferrua (2020, p. 2639 ff.).

¹⁰ Consider, for example, the “reasonable” or “realistic prospect of conviction” standard applied in various jurisdictions (England, Canada and, following a recent reform, in a very similar form in Italy) at the moment when the prosecutor has to decide whether or not to charge. On this point, see Della Torre (2024, p. 19).

¹¹ Consider, for example, the standard of “serious indications of guilt” required by Article 273 of the Italian Code of criminal procedure for the application of a precautionary measure.

¹² This is the case with BARD, moral certainty and *intime conviction*, which will be the focus of this study.

The recognition that there may be different standards of proof leads to the question of where one should be placed. Traditionally, the answer to this question has been that the level of these rules depends on the severity of the legal consequences of adopting a particular procedural choice. And it is precisely because rights and interests of fundamental importance to the accused are at risk of being compromised in the criminal sphere that the standards to be attained for the judgment at the end of criminal proceedings have historically been considered the most demanding compared to those used in other branches of law and, in particular, compared to those used in civil proceedings¹³.

This draws attention to another important function of such rules: their contribution to the distribution of the risk of erroneous judicial decisions¹⁴. To understand this better, it is useful to make a preliminary clarification. We can call: a) “false negative”, a decision that considers a true hypothesis unproven; b) “false positive”, one that considers a false hypothesis proven. Now, it is easy to see that as one decides to raise the level of a standard of proof, it becomes physiologically more difficult to achieve it in practice. This has two consequences: on the one hand, it increases the risk of a “false negative” (i.e. of failing to sufficiently prove a hypothesis that is in fact true) and, on the other hand, it reduces the risk of a “false positive” (i.e. of believing that a false hypothesis has been proved). Bearing this feature in mind, it is useful to understand, from another perspective, why the standard of proof set for the final decision in criminal proceedings is configured as the most demanding. This is because, with this type of rule, the systems do not seek so much to minimise the total number of errors in the reconstruction of the facts, but rather to limit the risk of occurrence of the category of error considered to be the most serious: the conviction of an innocent person (false positive)¹⁵. And it is clear that—as we shall see below—this asymmetrical nature of the standard of proof for a criminal conviction represents one of the law’s most important commitments to the presumption of innocence (Roberts and Zuckerman, 2022, p. 275 ff.)¹⁶.

¹³ For a recent study on this aspect, see Poli (2023).

¹⁴ It should be recalled that there is a complex debate about what exactly is distributed by standards of proof: is it the errors or the risk of errors? For a complete reconstruction of this debate, see Ferrer Beltrán (2021b, p. 115 ff.). In my view, the second of these alternatives is preferable for a fundamental reason: it is not only the standard of proof that influences the likelihood of the correctness of the decision on the facts, but also many other procedural rules (such as those governing the admissibility, gathering and evaluation of evidence). If this is the case, then logically the setting of a more or less demanding standard of proof has no effect (except very indirectly) on the reduction of the overall number of errors. Rather, they directly distribute the risk of error among the parties.

¹⁵ See Broun *et al.* (2014, p. 724) and Taruffo (2005, p. 117). For this very reason, the argument that a high standard of proof would help to reduce the overall number of factual errors is unconvincing. Indeed, this observation overlooks the fact that even wrongful acquittals are, strictly speaking, errors. In essence, setting high standards only leads us to favour one type of risk of error over the other. On this point, see Ferrer Beltrán and Tuzet (2018, p. 458).

¹⁶ As Iacoviello (2023, p. 301) reminds us, the presumption of innocence fixes the epistemological status of the hypothesis of innocence, stating: at the moment when the decision phase begins, the hy-

If this is true, the next step is to ask how many false acquittals a system is willing to tolerate in order to avoid a wrongful conviction. This is a question that has received mixed answers over the centuries. One of the best-known positions in this regard is that of William Blackstone, who stated: «the law holds [...] that it is better that ten guilty persons should escape, than that one innocent should suffer» (Blackstone, 1770, p. 352) But others have gone further, arguing that it's better to acquit more guilty people than to condemn an innocent one (Ferrer Beltrán, 2021b, p. 126). Beyond the rhetoric, it is clear that behind these divergences are different visions, influenced by political and moral choices, of the threshold at which we must set the level of evidence required for a criminal conviction.

It should be noted, however, that if one were concerned solely with the objective of protecting the innocent, one would be inclined to set a standard of absolute certainty for the imposition of a sentence. However, such a rule would create serious problems for the functioning of criminal justice systems (Ferrer Beltrán and Tuzet, 2018, p. 460). The most important of these is the fact that, given the nature of evidential reasoning, which is not logically deductive, but inductive and abductive, and therefore subject to margins of error (Taruffo, 1992, p. 166 ss.), such a decision rule «would result in acquittals in all cases, since there is no case without doubt» (Wharton, 1880, p. 2). It follows that such a policy option would only exponentially increase the number of wrongful acquittals, thereby preventing criminal justice systems from achieving two of their other inalienable objectives: ensuring the effective repression of crime and, consequently, the adequate protection of the community and victims from criminal behaviour. It is therefore no coincidence that, at the end of the 18th century, the eminent jurist Gaetano Filangieri warned against such a choice, stating that «five degrees more security in the courts would cost a hundred degrees less security in society» (Filangieri, 1806, p. 157).

Once we recognise the functions entrusted to standards of proof, we can see how crucial this type of legal rule is for the overall good functioning of a procedural legal system, and how important it is that they are designed to be easily applied in practice. Indeed, it is clear that the more the parties and the adjudicators are able to understand clearly whether or not the individual threshold of sufficiency of the proof has been reached, the better the functions assigned to the decision-making criteria will be fulfilled. But there is a serious impasse. There is no agreement on the best strategy to follow to create truly “efficient” standards of proof, i.e. capable of fulfilling the purposes assigned to them in an appropriate manner. In this respect, there is indeed a “great divide” between “subjective” and “objective” conceptions of standards of proof (Della Torre, 2015, p. 387 ff.; Haack, 2014, p. 52 ff.).

pothesis of innocence and the hypothesis of guilt are not on the same level, because the hypothesis of innocence has a privileged status. In fact, unlike the guilty hypothesis, the innocent hypothesis enjoys an initial plausibility derived directly from the law, without the need for proof.

Among the adherents of the first category are those who believe that it is possible to construct standards of proof as norms indicating the “*degree of persuasion*”¹⁷ or “*degree of conviction*”¹⁸ that the evidence must produce in the mind of the fact-finder in order for him to regard a hypothesis as proven. Essentially, according to the proponents of this interpretation, which, as we shall see, has been dominant in recent centuries, they should be constructed by referring to a graduated scale of individual belief, determined by the evidence, which rises progressively from the lower levels—complete uncertainty, suspicion and doubt—to the highest degree of conviction, which is never really attainable: absolute certainty as to the truth of a hypothesis¹⁹.

In recent decades, however, many authors, both in common law and civil law systems, have severely criticised the conception of these rules as “*standards of persuasion*”²⁰. In particular, it has been pointed out that the standards of proof, if conceived in this way, would be incapable both of adequately limiting the arbitrariness of the decision-makers and of enabling the parties to know with sufficient reliability when they would be able to obtain a decision in their favour. This is all the more true as the standards, if based on a certain level of conviction to be achieved, would end up depending on an often unpredictable fact: the ability of the adjudicator to be more or less easily persuaded by the evidence. Ultimately, there would be a risk of an exponential multiplication of the range of standards applied in practice, depending on the sensitivity of the fact-finder. As will be seen in more detail below, heterogeneous approaches have been taken to address this problem. On the one hand, there are those who have tried to interpret traditional rules in a more “objective” perspective, linking them to epistemic parameters rather than to the more or less fixed beliefs of the adjudicator (§ 7). A more radical approach, however, was to propose the total replacement of the classical standards by new ones, purified of references to subjective beliefs (§8). Whichever of these options is chosen, the common effect is to emphasise a more “objective” approach to standards of proof. That is, to see them as rules based not on a given degree of belief, but on a “*degree of rational justification*”, or a “*degree of warrant*” (Haack,

¹⁷ See, for instance, the position assumed by Bentham (1827, p. 71 ff.); Clermont (2013, p. 4); Broun et. al. (2014, p. 724 s.) and Wigmore (1940, p. 325). On closer inspection, it is precisely in the light of this subjective conception of the phenomenon of evidence that the terms “burden of persuasion” or “standard of persuasion”, which are prevalent in common law countries, could be explained.

¹⁸ As we shall see in § 5, this view has become widespread in continental Europe, mainly due to the success of the subjective standard of *intime conviction* of French origin.

¹⁹ As Haack (2014, pp. 48 ff.) recalls, in *In re Winship*, 397 U.S. 370 (1970), Justice Harlan wrote that the function of the standard of proof is to instruct the factfinder as to the “*degree of confidence*” our society thinks it should have in the correctness of factual conclusions for a particular type of adjudication. But for a similar construction, already see Bartolo da Sassoferrato (1562, Comment on Book XII of the Digest., tit. II, *lex XXXI*, f. 564, § 20 et ff.

²⁰ See Ferrer Beltrán (2021a, p. 40 ff.); Id. (2021b, *passim*); Haack (2014, p. 48 ff.); Laudan (2005, p. 95 ff.); Taruffo (2009a, 310).

2014, p. 56 ff.), that evidence must provide for a hypothesis about a fact in order for it to be accepted as proven²¹.

Having reached this point, it is worth noting that the clash between more subjective and more objective notions of the standards of proof is not a contemporary novelty, but rather a historically unresolved issue. It is true that there have been periods in time when one or the other of these views has prevailed. To understand this, it is essential to adopt a diachronic-descriptive approach, which aims to examine more closely the evolution of the main standards of proof used in Western legal systems to support a criminal conviction.

3. THE (SHATTERED) DREAM OF PRE-MODERN ROMAN-CANON SYSTEMS: THE RATIONALISATION OF STANDARDS OF PROOF THROUGH LEGAL PROOF RULES

3.1. The standard of *luce meridiana clariores* and its functional connection with the legal proof rules

A rule of the Codex of Justinianus states that a public accusation should not be made unless it was based on «suitable testimony, clear documents, or unquestionable circumstantial evidence, clearer than light itself»²². This rule was an important point of reference for the Roman-canon systems of the late Middle Ages, which used it as a benchmark to determine the “*gradus probationis*” for a criminal conviction²³. The general rule, repeated by criminalists in European countries until the end of the *Ancien Régime*, was that, in order to condemn an accused person, his guilt had to appear “clearer than the midday sun” (“*luce meridiana clariores*”)²⁴. This formula was intended to indicate a situation in which the evidence was particularly strong

²¹ However, some authors have questioned the possibility of formulating truly objective standards of proof: see in this respect the positions of Dei Vecchi (2022, p. 337 ff.) and by González Lagier (2020, p. 79 ff.). For a defence of “objectivist” positions, see, instead, Ferrer Beltrán (2021b, p. 203 ff.) and Aguilera (2021, p. 403 ff.).

²² The reference goes to C. 4.19.25, which reads: «*sciunt cuncti accusatores eam se rem deferre debere in publicam notionem, quae munita sit testibus idoneis vel instructa apertissimis documentis vel indicis ad probationem indubitatis et luce clarioribus expedita*».

²³ See, for example, Bianchi (1554, f. 264, § 10), who says: «*alius gradus probationis, quando procedendum est ad sententiam, & condemnationem, nam tunc luce meridiana clariores requiruntur*». In contemporary literature, see Alessi Palazzolo (1979, p. 3-6); Daniele (2009, p. 72); Garlati (2004, p. 395 f.); Picinali (2022, p. 19); Rosoni (1995, p. 70 f.).

²⁴ See, for example, Baldo degli Ubaldi (1578, tit. *De Probatione*, lex XV, *Sciunt cuncti*, f. 45, § 1); Bossi, (1564, tit. *de Oppositionibus contra testes*, f. 449, § 90); Claro (1586, *Liber V. § Fin. pract. Crim.*, quae. XXIII, f. 251, § 3); Carpzov, (1684, *pars. III, quae. CXX*, f. 178, § 12-14); Deciani (1613, cap. XXXV, tit. *De probationibus*, f. 358 § 20); Poullain-Duparc (1771, p. 112, § 41); Farinacci (1616, *pars. III, quae. LXXXVI*, f. 78); Le Brun de La Rochette (1611, p. 109); Mascardi, (1585, *conclusio* 459, f. 284, § 2 and 14).

or, depending on the terminology used by the various authors, “*plena*”, “*indubitata*”, “*apertissima*” or “*liquidissima*” (Allard, 1868, p. 244).

Another important point of reference in this field, taken from the Roman law, was a rule of the Emperor Trajan, which is collected in the Digest of Justinian²⁵. The reference is made to the so-called *lex absentem*, which enshrined the principle of «*satius enim esse impunitum relinqui facinus nocentis quam innocentem damnari*» (it is better to let the guilty go free than to condemn the innocent)²⁶. The strong valorisation of this rule makes it possible to understand how even the Roman-canon systems felt the need to structure the criminal standard of proof for a criminal conviction in such a way as to favour false acquittals over false convictions²⁷.

But the similarities do not end there. Roman-canon jurists also used the phrase “*luce clarioribus*” to emphasise the higher evidential threshold required to convict a subject in the criminal sphere than in the civil sphere (Alessi Palazzolo, 1979, p. 4). Let us take, in this regard, as an example one of the most famous Italian jurists of the 14th century: Baldus de Ubaldis. Well, it’s no coincidence that he felt the need to devote space to this issue in his commentary on the constitution *sciant cuncti*, concluding that criminal evidence must surpass civil evidence, not only in terms of the method by which it is obtained, but also in terms of its greater “demonstrative” value²⁸. Baldus was developing a line of thought that had been used before and after him, and which was a constant for Roman-canon lawyers who believed that «*ubi maius est periculum ibi cautius est providendum*»²⁹. This suggests that continental medieval systems had already established a proportional relationship between the value at stake in a given type of proceeding and the more or less demanding nature of the standards of proof. In criminal justice, especially when it came to imposing “ordinary punishment” (*poena ordinaria*), i.e. the punishment prescribed by law or custom, usually consisting of severe corporal punishment or the death penalty (*poena sanguinis*), the general rule was that the standard of proof for conviction should be as high as possible: *luce meridiana clariores*³⁰.

Having established these premises, we can now focus our attention on what, for the present analysis, turns out to be the main point of the experience of the Roman-canon systems: the attempt to rationalise the standard of *luce meridiana clariores* by establishing a rich set of legal proof rules, i.e. rules determining *a priori* the value to be given to a piece of evidence if it has certain quantitative or qualitative characteristics³¹. Refer-

²⁵ See D 48.19.5.

²⁶ On this point, see Dezza (2013, p. 41).

²⁷ See de Coras (1568, obs. CXL, 373, § 11); Farinacci (1607, cons. XX, f. 100, § 10); Gail (1613, ob. CXL, f. 253, § 11), who wrote: «*in dubio melius est nocentem absolvere, quam innocentem codemnare*».

²⁸ Baldo degli Ubaldi (1578, tit. *De Probation.*, lex XV, *Sciant cuncti*, f. 45, § 8).

²⁹ See Bartolomeo da Saliceto (1578, tit. *De probationibus*, f. 134, § 1).

³⁰ On this point see Bentham (1827, p. 93 ff.).

³¹ According to Ambos (2023, p. 170) «a system of legal proof consists of formal rules of evidence which instruct the adjudicator to take a certain fact as evidence only if this finding is in compliance with

ence is made to the development of the more general Roman-canon “system of legal proof”³², conceived at a doctrinal level from the late Middle Ages³³, later crystallised in a number of important legal texts of the sixteenth century³⁴, and then maintained in the main European legal systems until the time of the codifications³⁵. It was a system of rules designed to predetermine the character and value to be attached to each piece of evidence according to a complex “hierarchical” scheme (Levy, 1939). It is worth remembering that one of the main purposes for which this model was created was precisely to act as a counterweight to the power of judges and, in this sense, to constitute a respectable attempt to remove arbitrariness from the assessment of facts³⁶.

At a structural level, legal proof rules presented a recurring pattern, consisting of two distinct and interrelated elements (Daniele, 2009, p. 72). These rules indicated, on the one hand, the numerical and/or qualitative characteristics that a given piece of evidence had to possess and, on the other hand, the evidential value to be attached to it if the judge considered that the conditions prescribed by the rule were met. The bridge between these two elements was formed by an inference predetermined at the legal level, based on the generalisation of elements considered, according to the logic of the time, to be a valid gnoseological basis for legal knowledge, such as religious principles or Roman fragments.

Let us take, for example, one of the most classic rules of legal proof: the two-eye-witnesses rule, i.e. the rule which states that the testimony of two direct, concurring and credible witnesses is sufficient to prove a certain assertion about a fact³⁷. Now, such a rule could be divided into two parts. The first part specifies the number and quality of the witnesses (e.g. that they should be more than one, that they should have directly witnessed the commission of the crime and that they should not be in one of the numerous situations of incapacity to testify established by the sources of the time). The second specifies the value to be attached to their testimony if all the conditions set out in the first part are met (i.e. the fact that two unimpeachable eyewitnesses constitute “full proof”). Finally, it should be noted that the normative

the respective rules». He also recalls that there can be both positive rules—which positively indicate to the judge the value to be given to a piece of evidence if certain conditions are met—and negative rules, which prohibit the judge from reaching a particular conclusion if it does not have certain characteristics.

³² As Damaška (2019, p. 92 ff.) recalls, it should not be overlooked that the system of legal proof included not only rules of evaluation but also rules of admissibility of evidence.

³³ On this point see Alessi Palazzolo (1979, p. 8).

³⁴ The articulated *corpus* of Roman-canon legal proof rules is codified, for example, in the *Constitutio criminalis carolina* of 1532, an English translation of which can be found in Langbein (1974, p. 259 ff.).

³⁵ See, on this point, Damaška (2019, p. 7), who recalls that in some continental countries the theory of legal proof persisted throughout the twentieth century.

³⁶ See, on this point, Damaška (2019, p. 25) and Nobili (1974, p. 108-110).

³⁷ See art. 67 of the *Constitutio criminalis carolina*. For a comprehensive analysis, see Damaška (2019, p. 59-65).

inference linking these elements was based both on a maxim contained in the Bible³⁸ and on indications contained in Roman sources³⁹.

Having said this, it is now possible to focus attention on the close relationship that existed between the standard of proof of the *lucē meridiana clariores* and the traditional legal proof rules. In this regard, it should be noted that this standard established a threshold that had to be reached in order to convict, without, however, specifying the characteristics and the quality that each piece of evidence had to have in order to reach such a result. Well, one of the main purposes for which Roman-canon legal proof rules were developed was precisely to create a regime of “legal certainty”⁴⁰ on this point. By creating a system of rules on the value of evidence, an attempt was made to determine *ex ante* which and how much evidence could lead to the situation of evidential clarity required by the constitution *sciānt cuncti*⁴¹. In other words, the aim of the positive criteria for the evaluation of evidence was to solve the problem of the indeterminacy of the threshold of evidence to be reached in order to convict by the «compass of the law»⁴². Let’s take again the example of the traditional rule of two-eyewitnesses: it was a rule about the quality and quantity of the testimony, and at the same time it set a threshold of full proof that was considered to meet the relative standard of the *lucē meridiana clariores*. This is useful for understanding how, in the Roman canonical system of fact-finding, the standard of proof for a criminal conviction and the set of legal rules of proof, although strictly speaking different rules, were inextricably linked at a functional level⁴³. Indeed, the latter served to understand when there was sufficient evidence for a conviction.

3.2. The operational difficulties of the Roman-canon system of legal proof: the rise of the confession and the problem of the *indicia indubitata*

What we have seen so far should not lead us to conclude that the Roman-canon system of proof worked «by virtue of a rigid automatism» (Nobili, 1974, p. 113). In this respect, important studies (Alessi Palazzolo, 1979, p. 7; Damaška, 2019, p. 113; Rosoni, 1995, p. 40) have indeed disproved the traditional idea, resulting from the criticism of the Enlightenment reformers of the 18th century, according to which the legal proof rules obliged the judge to convict independently of his own convictions⁴⁴.

³⁸ *Deuteronomy*, 19:15-17.

³⁹ The reference is to Constantine’s constitution, contained in C Th. 11.39.3.1 and repeated in C. 4.20.9.

⁴⁰ On this notion, widespread above all in nineteenth-century treatises, Mittermaier (1850, p. 100 s.).

⁴¹ See de Damhouder (1556, ch. 49, f. 144, § 6-8).

⁴² The expression comes from Arrivabene (1814, p. 24).

⁴³ On the functional connection between standards of proof and legal proof rules, see Tuzet (2021, p. 107).

⁴⁴ See Touret, speech of the 11.1.1791, in Mavidal and Laurent (eds.) (1885, p. 132). In this respect, Esmein’s (1882, p. 260) image of the Roman canonical judge as «un clavier qui répond inévitablement lorsqu’on frappe certaines touches» remains famous.

More specifically, these studies have argued that the normative inference underlying the legal proof rules was only triggered when judges considered that the evidence obtained met several requirements, both formal and substantive, which ultimately left them with a not inconsiderable margin of discretion (“*arbitrium*”)⁴⁵.

In so far as it is relevant here, it should be noted that there is one aspect in which the Roman canonical system was characterised by a certain rigidity: the attempt to establish *ex ante* what and how much evidence was necessary to condemn the accused to severe corporal or capital punishment (*poena sanguinis*). This recalls the famous division between the standard of “full proof” (*probatio plena*), i.e. that which «*tantum fidem facit, quantum ad finiendum controversiam sufficit*» (Mascardi, 1585, quae. IV, f. 4, § 16) and the standard of “half proof” (*probatio semiplena*)⁴⁶, i.e. that which remains incomplete⁴⁷. This distinction had important legal consequences: only the former were considered equivalent to the standard of the *luce meridiana clariores* and thus justified the application of the most severe criminal sanctions (Alessi Palazzolo, 1979, p. 43).

It should be noted, however, that the Roman-canon system tended to limit the qualification of “*plena probatio*” to a limited range of types of evidence⁴⁸, capable of directly establishing the fact of the offence, such as the rule of two eyewitnesses, or judicial confession⁴⁹. By contrast, direct evidence that did not meet certain requirements, or evidence that was only circumstantial, was classified as “weak”⁵⁰. For this reason, it was classified as “half proof”⁵¹ or in an even lower category⁵² and was generally considered inappropriate as a basis for applying the most severe penalties⁵³.

However, it is not difficult to see that this hierarchical construction could give rise to considerable practical problems. Indeed, since it is not always easy to obtain evidence that meets all the requirements of full proof, such a system, if rigidly understood, would risk preventing the application of ordinary punishment in a significant number of cases, even in the presence of serious evidence of guilt⁵⁴. With all that this would entail in terms of a loss of the repressive capacity of penal systems and an

⁴⁵ See Damaška (2019, p. 59-81). See, in support of this thesis, the words of Farinacci (1607, *concl.* 70, f. 336): «*probationes omnes sunt iudici arbitrariae*».

⁴⁶ On this subject, see Mascardi (1585, quae. IV, f. 4, § 17), who defines the half-proof as follows: «*per quam rei gestae fides aliqua fit iudici, non tamen tanta, ut iure eam debeat sequi in sententia dicenda*».

⁴⁷ On this distinction, see Rosoni (1995, p. 73 ff.).

⁴⁸ On this point see the authoritative opinion of de Damhouder (1556, cap. 49, f. 143 f.), who included in this category only the cases of two-eyewitnesses, judicial confession and *evidentia facti*.

⁴⁹ For the characteristics that the confession had to have to be classified as full proof, see Damaška (2019, p. 69 ff.).

⁵⁰ See again de Damhouder (1556, cap. 49, f. 144, § 4-7).

⁵¹ See Rosoni (1995, p. 81).

⁵² See, for example, Claro (1586, quae. XX, f. 240, § 1): «*et hoc credo proprie in hac materia appellari posse indicium, quod silicet sit plus quam simplex preasumptio, et minus quam semiplena probatio*».

⁵³ On this subject, see Alessi Palazzolo (1979, p. 14); Rosoni (1995, p. 124 ff.).

⁵⁴ See, in this regard, Alessi Palazzolo (1979, p. 55) and Rosoni (1995, p. 81).

increased risk of the propagation of revenge spirals by victims or their families. And it was also the desire to limit these risks that led the Roman-canon systems to rely more and more on a category of full proof that was considered more easily available than the others: the confession, which in time became the *regina probationum*. As is well known, the desire to obtain many confessions led to a cascade of poisoned fruits. In particular, it contributed to the spread of that unacceptable method of extracting statements against one's will which was used, with varying degrees of success, by continental systems until the Enlightenment: judicial torture⁵⁵.

To these problems were then added other delicate theoretical questions arising from the difficulty of reconciling the above-mentioned devaluation of circumstantial evidence with some opposite indications coming from Roman law (Alessi Palazzolo, 1979, p. 55). In this regard, it is sufficient to recall once again the constitution *sciunt cuncti*, which, as we have seen, not only attributed to circumstantial evidence, provided that it was unquestionable circumstantial evidence (*indicia indubitata*), the qualification of *luce clarioribus* evidence, but also equated it with two categories of full proof, namely testimony and documentary evidence. This obviously made it difficult to justify the tendency to deny even decisive value to this category, since it was valued by the same source from which the main criminal law standard was derived. In view of this, it is not surprising that the question of the role of *indicia indubitata* was one of the most debated issues in Europe up to the time of the codifications⁵⁶.

One of the first problems was the definition of the concept. It was a question that was as crucial as it was difficult to resolve⁵⁷, partly because the attempt to settle it by listing all the typical hypotheses of unquestionable circumstantial evidence did not provide satisfactory answers (Alessi Palazzolo, 1979, p. 56). As we shall see in the next paragraph, the discussion on this point was enriched by new positions in the 17th and 18th centuries, when an attempt was made to solve the problem by using the concepts—important from late medieval and modern theological reflection—of moral certainty and reasonable doubt. Whatever the case, one fact is clear: as the writers of the time recognised, the decision as to whether or not to classify circumstantial evidence as “*indubitata*” could ultimately only depend on the decision of the individual adjudicator⁵⁸. In the end, this gave the judges a further (and even more important) margin of discretion in the assessment of evidence, capable of making the Roman-canon system much more flexible than it might appear in the abstract.

⁵⁵ See Mittermaier (1859, p. 291), who recalled that «the more widespread the use of torture became, the more the value of the confession increased». On this point, see Fiorelli (2023).

⁵⁶ For a detailed reconstruction, see Raynaldi (1735, cap. XXI, § IV & V, f. 259, § 27 ff.). In contemporary literature, see Rosoni (1995, p. 135 ff.).

⁵⁷ It is no coincidence that Raynaldi (1735, cap. XXI, § IV & V, f. 260, § 32) remarked that this question «*diversimode a Doctoribus definiuntur*».

⁵⁸ See Basilio (1691, f. 2) who stated that «*quae sint indicia indubitata melius arbitrio iudicis relinquendum*».

The above consideration is even more evident in the light of the debate that has developed as to whether or not it is possible to convict in criminal cases on the basis of this type of evidence⁵⁹. A representative picture of the heterogeneous answers given to this question over time can be drawn by an author belonging to the mature period of Italian criminalistics: Giulio Claro. After admitting that «*in hoc articulo Doctores multum varie loquuntur*» (Claro, 1586, Liber V. § *Fin. pract. Crim., quae*. XX, f. 241, § 6), he summarised the main positions that had developed on the subject up to the 16th century. A first strand of interpretation, Claro recalls, proved to be particularly conservative: it took the general devaluation of circumstantial evidence described above to its extreme, concluding that even in the presence of genuine, *indicia indubitata*, no criminal conviction could be obtained. A second group of authors, however, took the opposite view. Relying on the equivalence established by the *sciunt cuncti* constitution between the *indicia indubitata* and the standard of *luce meridiana clariores*, the exponents of this current of thought considered that this category could justify the imposition of the ordinary penalty, even if it was the death penalty. It should be noted that this type of approach has been particularly successful where certain types of crime have been prosecuted, including those which are judged to be “atrocious”, “hidden” or “difficult to prove”⁶⁰.

Claro’s work is also useful because it allows us to see that, among the extreme opinions we have described, other, more middle-of-the-road positions emerged and gradually prevailed, which were able to find more flexible solutions. The awareness that the Roman-canon system could not afford to abstain from conviction even on the basis of circumstantial evidence led to the search for a way out of the narrow dichotomy between acquittal on the one hand and the application of ordinary punishment on the other. To this end, a strategy was planned based on the following idea: if the evidence gathered during the trial did not allow the standard of “full proof” to be reached, but elements suitable for reaching the threshold of *indicia indubitata* had been gathered, a lesser punishment, chosen at the discretion of the judge, could be applied instead of the death penalty⁶¹. This compromise solution became known as *poena extraordinaria* (extraordinary punishment)⁶². A phenomenon that began in the Middle Ages and crystallised over the following centuries, reaching European proportions at its height.

What we have just seen shows how continental jurists responded to the practical difficulties caused by the rigidity of the Roman-canon system of proof, not only by insisting on confession, but also by establishing a proportional relationship between the degree of proof and the punishment to be inflicted (Alessi Palazzolo,

⁵⁹ On this point, see Rosoni (1995, p. 138 ff.).

⁶⁰ See Rosoni (1995, p. 140).

⁶¹ Claro (1586, *Liber V. § Fin. pract. Crim., quae*. XX, f. 241, § 6) where he stated: «*alij dicunt, quod ubi ex preasumptionibus proceditur multum debet iudex temperare suam sententiam, & maxime ne condemnet nisi raro, & modice ad poenam*».

⁶² In contemporary literature, on this point see Alessi (1979, p. 19 ff.); Damaška (2019, p. 106-111); Picinali (2022, p. 20).

1979, p. 21), which ultimately determined a substantial doubling of the standard of proof for conviction. If it is true that the general standard remained formally that of the *luce meridiana clariores*, it is also true that in practice it was divided into two quantitative and qualitative thresholds which, for the same offence committed, led to different results in terms of punishment. The optimal threshold was reached only when full proof was obtained; and since it was considered that such evidence could establish a situation of “legal certainty” as to the guilt of the accused, the imposition of the most severe penalties, in particular the death penalty, was considered justified. The second threshold was that of *indicia indubitata*: this category, which presupposed the judge’s belief in the guilt of the accused, but not the existence of evidence that could be classified as full proof, was considered less reliable and justified only the imposition of penalties milder than the death penalty. It is useful to note, moreover, that when the Roman-canon system sought to prohibit the imposition of the *poena sanguinis* in the absence of full proof, it proved to be more protective from this point of view than those contemporary legal systems which still allowed the death penalty to be applied on the basis of the BARD standard; that is, a criterion which, as we shall see shortly, is at a level comparable to that of the *indicia indubitata*⁶³.

The problem was that the standard of “full proof” turned out to be so demanding that it was often impossible to meet in practice, which favoured the development of creative solutions, one of the most important of which was precisely that of extraordinary punishments⁶⁴. Although this stratagem had its advantages, both practical and religious⁶⁵, it was not without its costs. Such an operation further increased the power of the professional judiciary body, both at the procedural level and in the overall assessment of the evidence and, above all, at the level of the choice of the sanctions. This helps us to understand how the attempt of the Roman-canon system of fact-finding to make the standard of proof for a criminal conviction objective, by establishing a rigid set of evidence assessment criteria, proved to be a failure long before its final crisis in the Enlightenment. Over time, the mechanism of legal proof was no longer able to satisfy the need for “legal certainty” that was its *raison d’être*, confirming the ancient rescript of the Emperor Hadrian, according to which «*quae argumenta ad quem modum probandae cuique rei sufficient, nullo certo modo satis definiri potest*»⁶⁶. It is not difficult to see how the gradual realisation of this ultimately favoured the implosion of the system, from an epistemological rather than a moral point of view. It should be noted that as the demands for exemplarity and repressiveness typical of the states of the *Ancient Régime* grew, some legal systems tried to

⁶³ This is pointed out by Damaška (2019, p. 108).

⁶⁴ For further examples, see Damaška (2019, p. 111 ff.), who also includes in this category the so-called “intermediate judgments”, on which see Picinali (2022).

⁶⁵ In effect, this instrument protected judges from the ethical-religious responsibility of pronouncing a death sentence on the basis of a decision that depended entirely on their inner convictions. This was a very significant advantage, since taking the life of an innocent person risked becoming a mortal sin. For more detailed reflections on this topic, see Damaška (2019, p. 91) and Whitman (2008, p. 3).

⁶⁶ The passage is included in D. 22.5.3.2. In curiously similar terms, see Ambos (2023, pp. 177 f.).

find ways of limiting the practice of extraordinary punishment. And it is precisely in the context of these authoritarian reactions that the concepts of moral certainty and BARD began to take their first steps in the legal field, at least for some civil law systems. However, instead of revitalising the traditional system of legal proof, these experiments contributed to its final collapse.

4. THE ORIGIN OF MORAL CERTAINTY AND REASONABLE DOUBT

4.1. The theological background

As is well known, in the Middle Ages there was a widespread belief that in human affairs, unlike in logic or mathematics, the absolute certainty of the truth of a proposition was unattainable⁶⁷. In the wake of this awareness, the notion of certainty was conceived as a graduated scale: one of the most important classifications «was a sequence of three steps [...] that led down from mathematical (or metaphysical) over physical to practical certainty» (Shüssler, 2009, p. 453)⁶⁸. As far as is relevant here, it should be noted that such a tripartite division had a decisive influence on the development of the standards of moral certainty and BARD, the origins of which will be examined in this section.

The first element to be clarified is the fact that these criteria do not originate in the world of law, but in the Christian theological-moral field. To understand this, it is useful to recall that the paternity of the concept of “moral certainty”—from which the BARD has been derived over the centuries—is attributed to the French theologian Jean Gerson (Franklin, 2001, p. 69), who lived at the time of the great Western Schism at the end of the 14th century. He developed this concept to address one of the most important concerns of societies at the time, which were particularly troubled by moral doubts due to the conflict between the Roman and Avignon papacy: to determine when a person could be said to be sufficiently certain to act without committing a sin (Schüssler, 2009, p. 453). In terms of content, even in Gerson’s original construction, moral certainty would have implied only a high probability of the truth of a hypothesis⁶⁹. Consequently, this standard would have allowed for the theoretical possibility of the contrary and would have been attainable even if some (unspecified) residual scruple of uncertainty had remained in the human soul⁷⁰. Nevertheless, such a standard would still have been an indication of a sufficiently firm conviction for the individual’s salvation from sin⁷¹. It is not difficult to see that in

⁶⁷ See Thomas Aquinas (1265-1274), *Summa Theologiae*, II- II, 70, art. 2, n. 1, who concluded that in human affairs there can be no “demonstrative” certainty, but only a “*probabilis certitudo*”.

⁶⁸ An updated version of this tripartite division appears in Baron (2022).

⁶⁹ See, on this point, Jerson (1706, *De consolatione theologiae*, lib. IV, *prosa secunda*, p. 171 f.).

⁷⁰ See Biel (1574, *distinctio XVI, quaest.* III, f. 515).

⁷¹ See Jerson (1706, *Gersoniana*, lib. IV, cap. III *De Morum Disciplina*, CXLIX).

choosing such a rule to indicate «a very high but not complete degree of persuasion» (Franklin, 2001, p. 69), Gerson was trying to strike a balance between widespread scrupulousness in moral matters on the one hand, and vigorous action on the other.

Thanks in part to its flexibility, the concept of moral certainty was an immediate success. Used by some of the most influential thinkers of the late Middle Ages⁷², this idea was able to spread as early as the fifteenth century and, in the course of modernity, became an indispensable point of reference for all those who wanted to answer the question of what faith was necessary in human affairs in order not to risk making a wrong decision⁷³.

After the Lutheran Reformation, efforts to clarify its meaning were intensified⁷⁴. From this point of view, an important contribution came from the clash between “casuistry” theologians that took place between the 16th and 18th centuries⁷⁵. In fact, it was in this context that one began to describe as “reasonable” or “rational” the hesitation that could prevent one from arriving at moral certainty⁷⁶. This makes it possible to understand how, originally, the standard—subjective because it refers to a state of mind—to be reached in resolving cases of conscience was that of moral certainty, which could only be said to be satisfied when a subject succeeded in eliminating all reasonable hesitation about the goodness of a solution (Shapiro, 1991, p. 21). From that time on, such theories were constantly used, even to the point of being endorsed by the leadership of the Catholic Church (§ 4.2).

However, the concept of moral certainty, understood as the highest possible certainty about past events, did not remain confined to the theological sphere, but quickly spread to many other fields of knowledge⁷⁷, including law⁷⁸. This is hardly surprising, as it was a very valuable criterion in a context where the fact-finders had to make decisions about past events that were extremely delicate from both a legal and an ethical-religious point of view. It is therefore easy to understand why it was decided to transfer this criterion to criminal justice. On the one hand, the opportunity was seized to use as a decision-making rule a standard capable of protecting the individual from the risk of a wrongful conviction, in accordance with the old adage

⁷² See Antonino da Firenze (1740, p. 204).

⁷³ See Shüssler (2009, p. 453).

⁷⁴ See Suarez (1741, p. 319), who said «*certitudo moralis omnem propriam dubitationem excludit*».

⁷⁵ On this point, see Shapiro (1991, p. 13 ff.) and Shüssler (2019, p. 60 ff.).

⁷⁶ See, among many others, Angles (1587, *concl.V*, f. 73); Concina (1761, p. 25); Caramuel Lobkowitz (1675, p. 255); de Lugo (1696, p. 72); Martinon (1663, p. 263); Medina (1580, p. 624); Patuzzi (1790, p. 291); Sanchez, (1622, p. 405); Verricelli (1653, p. 66). For the English experience, see Smith (1748, p. 42), who explicitly linked “moral certainty” with the absence of any “reasonable doubt”.

⁷⁷ For example, the concept of moral certainty was used by some of the fathers of modern scientific thought, such as Descartes (1667, p. 483), who defined such certainty as that which «*est aussi grande que celle dont nous n'avons pas coutume de douter touchant la conduite de la vie*». For many other examples, see Shapiro (2022, p. 127 ff.).

⁷⁸ See Shapiro (2009, p. 268 f.) and Ead. (2012, p. 19 f.).

of *in dubio pro reo*, but not so high, given its probabilistic nature, as to sacrifice the repression of crime. On the other hand, an instrument has been introduced—approved by the most authoritative theologians as the best criterion for resolving cases of conscience—which could relieve the adjudicators of the fear of committing a mortal sin, given the ever-present risk of condemning an innocent person (Whitman, 2008). It is worth noting that this process of introducing moral certainty and reasonable doubt into the field of law took place in the late modern period, both in the accusatory systems of the common law and in the galaxy of inquisitorial systems of the Roman-canon tradition. Let us see how it happened.

4.2. The Anglo-Saxon experience

As far as the Anglo-Saxon legal systems are concerned, the historical development of the modern standard of proof for a criminal conviction has been the subject of extensive studies, the main results of which are worth mentioning here⁷⁹.

In this regard, the common law literature has highlighted how, until the 15th century, juries were given summary instructions which merely emphasised the need to decide «according to the evidence and your conscience» (Shapiro, 2009, p. 262). However, it was not until the 16th and 17th centuries that the need was felt to specify the threshold that had to be reached for an offence to be considered proven. The English jurist Matthew Hale, for example, worked in this direction, and his works helped to consolidate concepts that were already widespread in theological debate, such as the idea that in doubtful cases one should avoid acting and take the decision most favourable to the accused⁸⁰. And it was in this context that, through a process of gradual linguistic refinement that did not affect the substance of the norm, it began to be affirmed that the triers of fact could not convict unless they had a “satisfied conscience”, a “satisfied understanding” or even a “moral certainty” of the guilt of the accused (Shapiro, 2014, p. 261 ff.).

It should be added, however, that scholars agree that the last formula to appear in court would have been that of BARD, which only began to be used in the 18th century⁸¹. Suffice it to say that the first documented practical use of reasonable doubt in the Anglo-American context can traditionally be traced back to the Boston Massacre trial of 1770⁸². In that trial, Robert Treat Paine addressed the jury in the following terms: «if therefor in the examination of this Cause the evidence is not sufficient to convince you beyond reasonable doubt of the guilt of all or of any of the prisoners

⁷⁹ See Langbein (2003, p. 261 ff.); Morano (1975, p. 507 ff.); Shapiro (1991); Ead (2012, p. 19 ff.); Whitman (2008).

⁸⁰ See Hale (1778, p. 300).

⁸¹ On this point, see Langbein (2003, p. 261 ff.); Shapiro (2009, p. 274 f.); Whitman (2008, p. 192 ff.).

⁸² See Morano (1975, p. 507); Shapiro (2014, p. 23); Whitman (2008, p. 193).

by the benignity and reason of the law you will acquit them. But if the evidence be sufficient to convince you of the guilt of all or of any of the Prisoners by the benignity and reason of the law will require you to declare them guilty»⁸³. On this point, however, it should be noted that there is nothing in the record of the Boston cases to suggest that the reference to the standard was intended by the parties as a particular novelty; this is hardly surprising, precisely because the use of similar phrases «was consistent with the notions of “belief”, “satisfied conscience” and “moral certainty” used in and out of the courtroom» (Shapiro, 2014, p. 22).

Finally, as Langbein has shown, at the end of the eighteenth century the BARD standard had not yet become a rule of law in the sense that it was invariably applied, which is why we still find different judges formulating the standard of proof differently (Langbein, 2003, p. 264). Indeed, an analysis of the case law of the time shows that there were cases in which such a rule of law had not yet been established, or in which reference was still made to criteria that were not capable of satisfying the “beyond reasonable doubt” requirement. As we shall see in § 6, it was only later, thanks in particular to the treatises and case law of the 19th century, that the BARD rule became firmly established in Anglo-Saxon theory and practice.

4.3. The Roman-canon experience

In continental legal systems, the origins of modern standards of proof for a criminal conviction are more complicated and therefore need to be reconstructed in more detail. First of all, it should be remembered that the concept of moral certainty was disseminated in the culture of the *ius commune* not only through the numerous legal studies written by members of the clergy, but also through the works of a number of great philosophers, including the great German jurist Samuel Pufendorf. In his *De Jure Naturae et Gentium*, he adapted the concept of *certitudo moralis* to the legal context, defining it as a «*firma [...] presumptio, valde probabilibus rationibus subnixta, & quae nisi rarissime fallere possit*» (Pufendorf, 1759, p. 27 § 11).

These concepts were soon applied in practice to solve complex problems of Roman-canon criminal justice and especially those concerning the role of the *indicia indubitata*. The experience of some late modern Italian jurisdictions is particularly instructive in this respect. Faced with a particular crime wave, some countries reduced the *gradus probationis* for the imposition of a sentence to an “ordinary penalty” under the traditional standard of full proof. They did so by enacting a series of exceptional rules that empowered judges to impose the *poena sanguinis* on the basis of the lower standard of unquestionable circumstantial evidence. And precisely to compensate for

⁸³ The quotation is from the book *The Trial of the British Soldiers*, William Emmons (1824, p. 118). Scholarship is divided as to whether or not the use of the BARD in this case resulted in a lowering of the standard of proof compared to previous practice: for a summary of the various positions, see Sheppard (2003, p. 1191 f.).

this lowering of the standard of proof, reference was made not only to the concept of moral certainty, but also to criteria very similar to those of the BARD.

The 17th century legislation of the Kingdom of Naples is one of the earliest examples. On 20 September 1621, Viceroy Zapata issued Pragmatics XII, in which he not only ordered the courts to «follow the opinion of giving the ordinary penalty for unquestionable circumstantial evidence», but also expressly stated that this concept should be understood as evidence capable of «inducing the mind of the judge to firmly believe that the crime has been committed by the suspect, calming his intellect in this firm belief»⁸⁴. As could be expected, this reform raised a new problem: understanding what “firmly believe” means⁸⁵. As the jurist Tommaso Briganti recalls⁸⁶, there were many answers to this question, including two of great interest for the present analysis. Firstly, it is worth noting the opinion of those who considered that the standard could only be reached if the judge was able to exclude all “actual” (i.e. arising from the evidence obtained *ex actis*) and, in some cases, even “virtual” (i.e. even merely theoretical) doubts as to the defendant’s innocence⁸⁷. A second group of authors, led by Carlo Antonio de Rosa, echoing the positions of several theologians of the time, drew a less demanding standard from the law, arguing that the concept of “firm belief” should be understood precisely as a synonym for “moral certainty”⁸⁸.

In the 18th century, the debate spread to other parts of the Italian peninsula. This was the case, for example, in the Church State, where Pope Benedict XIV, noting an increase in certain types of crime, issued a chirograph on 2 January 1743, not only authorising his courts to sentence certain crimes to death, even on the basis of *indicia indubitata*, but also identifying with this category circumstantial evidence that «removes from the mind of the judge any reasonable hesitation that the crime might otherwise have been committed»⁸⁹. Given the above theological and moral origins of reasonable doubt, it is not surprising that Pope Benedict XIV explicitly codified this standard in his legislation. In doing so, the Pope showed himself to be in agreement with what some of Christianity’s most authoritative thinkers had affirmed in their works. But there is more. Given the close link that has always existed between the BARD and the *certitudo moralis*, it is not surprising that this expression too was soon used to interpret Benedictine chirograph. In this regard, it may be useful to refer, for example, to the thought of the criminalist Giovan Battista Gallucci, who, a few years

⁸⁴ Prammatica XII, in *Pragmaticae, edicta, decreta regiaeque sanctiones regni Neapolitani*, t. II, Jacobi Raillard, 1682, p. 836. On the context in which this Neapolitan law was developed, see Alessi Palazzolo (1979, p. 192 ff.).

⁸⁵ It will not go unnoticed that the expressions used here are very similar to the English expression satisfied belief.

⁸⁶ See Briganti (1842, p. 152 s.).

⁸⁷ In this sense, see, for example, Rovito (1634, f. 312, § 9 f.).

⁸⁸ See De Rosa (1680, f. 18, § 32) and Briganti (1842, p. 152 s.).

⁸⁹ See Benedict XIV, *Bullarium*, t. I, Ioannis Gatti, 1778, p. 104. On this measure, see *amplius* Ala (1829, p. 214-221).

later, not only reflected on the concept of moral certainty, but also argued that such a state of mind could only be reached when the judge was able to exclude *omnem rationabilem haesitationem* regarding guilt⁹⁰. Moreover, it should not be forgotten that it was later the ecclesiastical authorities themselves who combined the two locutions in a single provision. This was the case, for example, in an Edict of 1805, in which the absence of any reasonable hesitation was expressly equated with «moral certainty as to the guilt of the accused»⁹¹.

Finally, the Grand Duchy of Tuscany has also adopted a law with a similar tenor. We are referring to article II of the law of 15 January 1744, which states that «in criminal cases the evidence resulting from *indicia indubitata* [...] from which a moral certainty arises against the accused [...] is and shall be full proof and sufficient for the punishment of death»⁹². This rule was so successful in Tuscany that the judiciary sought to reassert its power even after the approval of Grand Duke Peter Leopold's famous reform of the penal system in 1786⁹³, article 110 of which seemed to reinstate the prohibition of sentencing to death on the basis of circumstantial evidence⁹⁴.

This early codification of the standard of moral certainty and “*beyond any reasonable hesitation*” in some Italian states contributed to a general circulation of these concepts in the continental procedural literature of the time, to the point of making them, as we shall see, one of the workhorses of Enlightenment thinkers. It is equally undeniable, however, that this phenomenon, at least from the perspective of the standards of proof, marked a step backwards with respect to the classical structure of the Roman-canon systems of proof. In fact, as we have seen, the threshold of *indicia indubitata* was traditionally considered lower than that of full proof and, for this very reason, was usually not considered sufficient to justify the imposition of the *poena sanguinis*⁹⁵. Nevertheless, the difficulty of reaching the standard of full proof was such that, in order to meet the greater repressive demands of modern absolute regimes, a downward adjustment of the decision-making rules for a criminal conviction was imposed. In this way they were brought down to the admittedly lower but more realistic level at which they still stand today. In short, it can be said that rules such as those of Naples, Rome or Tuscany were an important step that helped the Roman-canon criminal justice system to make the definitive transition from a law of evidence based on “legal certainty” to one based on the “moral certainty” of the decision-maker, reached after a free evaluation of the evidence by the fact-finder (Nobili, 1974, p. 117-119).

⁹⁰ See Sinistrari (1754, p. 506 f., § 42 ff.).

⁹¹ See *Diario ordinario*, Rome 1805, no. 80, p. 3.

⁹² See *Ordini diversi da osservarsi nelle cause criminali. Pubblicati sotto dì 15 Genn. 1744*, Firenze 1745.

⁹³ Reference is made here to the *Legge criminale Toscana* (the so-called “Leopoldina”) of the 30th of November 1786.

⁹⁴ On this point, see Carmignani (1852, p. 354 ff.).

⁹⁵ It is for this very reason that Carmignani (1852, p. 293) criticised these rules.

Finally, it should be noted that the examples given so far refute the idea that the BARD standard is historically typical only of common law accusatory systems⁹⁶. As we have seen, from a historical perspective, it is a theological standard designed to better explain the concept of moral certainty, which entered the field of law in modern times. And it does not matter that, strictly speaking, the standard of “beyond any reasonable hesitation” was codified in a papal chirograph several decades before the earliest recorded occurrences of the BARD in Anglo-American jurisprudence; an element which in itself would end up attributing primacy in this respect to continental law. It is interesting to note that in the eighteenth century, despite the residual survival of the general system of legal proof on the European continent and the innumerable differences between the Anglo-Saxon accusatory model and the continental inquisitorial one, the problems of the standard of proof were in many ways the same, as were the lexicon, the positions and the solutions. This reinforces the idea that moral certainty and the BARD are concepts that have deep transversal roots in the culture—theological, philosophical and, later, legal—of the entire West⁹⁷.

5. THE COLLAPSE OF THE SYSTEM OF LEGAL PROOF AND THE EMERGENCE OF THE *INTIME CONVICTION*

5.1. The Enlightenment criticism of legal proof rules

During the 18th century, the reform of criminal procedure became one of the central concerns of the Enlightenment (Nobili, 1974, p. 117 ff.). The thinkers of the time identified the inquisitorial model of criminal justice as an instrument incompatible with the theory of the natural rights of the person, which had begun to develop in the 17th century. As a result, all the pillars of the Roman-canon criminal justice system were criticised, especially by the Italian and French Enlightenment⁹⁸. The written and secret procedure, the presence of professional judges with wide *ex officio* powers, torture, cruel punishment and, finally, the rules of legal proof were considered ruins to be overcome by the light of reason. On the contrary, the English accusatorial system, based on the opposite principles of orality, cross-examination, publicity, free evaluation of evidence and, above all, the presence of a jury trial, was seen as a model to be followed (Damaška, 2019, p. 118).

Insofar as it is of interest here, it is worth noting how, in this context, the standard of moral certainty also made a qualitative rise on the Continent, from being an exceptional rule relevant mainly to circumstantial evidence to being adopted as a general

⁹⁶ Even the great Italian jurist Cordero (2012, p. 995), called the BARD standard an “Americanism”.

⁹⁷ For a similar perspective, see Whitman (2008, p. 5).

⁹⁸ On the heterogeneous path taken by the German Enlightenment in this respect, see Ambos (2023, p. 171 ff.) and Nobili (1974, p. 135 ff.).

criterion «*nécessaire et suffisant pour émettre la décision judiciaire*» (Padoa-Schioppa, 1999, p. 122). In other words, moral certainty was conceived as the new ideal criminal justice standard for a criminal conviction, capable of indicating when evidence was suitable for achieving the still remembered ideal of the *luce meridiana clariores*.

It should be noted, however, that two factors in particular favoured the rise of the “moral criterion” and the gradual disintegration of the standard of full proof, associated with a set of legal proof rules: on the one hand, the Enlightenment’s renewed faith in “common sense” and human reason, and on the other, the great favour shown to the institution of the jury, understood as the “palladium of liberty”⁹⁹. In other words, the Enlightenment thinkers wanted to replace the professional judges linked to central power, who decided by applying the complex network of Roman-canon rules for evaluating evidence, with ordinary citizens who were called upon to decide, according to their common sense, whether the evidence was sufficient to establish in their minds a full “moral certainty” of the guilt of the accused. It is therefore clear that, in this perspective, moral certainty also took on the role of a criterion aimed at “democratising” the law of evidence¹⁰⁰.

The manifesto of this way of thinking is Cesare Beccaria’s *On Crimes and Punishments*. In this book, a special place is given to moral certainty, which is not only qualified as a high probability, but is also compared to that kind of conviction «that determines everyone in the most important actions of his life» (Beccaria, 1880, p. 135). These are, of course, expressions similar to those that have been used for centuries in the language of Christian theologians or by philosophers such as Descartes, traces of which, thanks to Beccaria’s worldwide success, still persist in some legal systems¹⁰¹.

Beccaria did not stop there, however, and drew a link between moral certainty and the institution of the jury. After admitting that «this moral certainty of proofs, however, is easier to feel than to define with exactitude», he described as excellent that law which provides for trial by jury, because there is «more safety in the ignorance which judges by sentiment than in the knowledge which judges by opinion» (Beccaria, 1880, p. 136). According to him, all that is needed to evaluate the evidence is «a simple and common good sense, a faculty which is less fallacious than the learning of a judge, accustomed as he is to wish to find men guilty» (Beccaria, 1880, p. 136). Clearly, the great Italian jurist was doing nothing more than taking the Enlightenment belief in reason and human judgement to its logical extreme.

Notwithstanding the breaches that moral certainty and reasonable doubt had opened up in the traditional structure of the Roman-canon model of fact-finding, the reversal of perspective thus theorised was nevertheless marked. Such was the distrust of the traditional system that it led to the idea of the need to abandon altogether the regime of “legal certainty” in matters of evidence in favour of one based

⁹⁹ See, on this point, Nobili (1974, p. 131).

¹⁰⁰ See Leclerc (1995, p. 206).

¹⁰¹ See, in this respect, following the note n. 131.

entirely on human “moral certainty”. Thus, on the European continent, there was a definitive shift from a system based primarily on “objective” standards of proof, because they were linked to a network of criteria for evaluating evidence, to one based entirely on standards linguistically conceived as degrees of persuasion, because they were linked to the stability of the decision-maker’s beliefs.

It should be noted, however, that some Enlightenment thinkers were aware of the side effects of such a complete reversal of perspective. In this regard, it is worth quoting Voltaire, who, while severely criticising the system of legal proof, concluded: «on serait tenté de souhaiter que toute loi fût abolie, et qu’il n’y en eût d’autres que le conscience et le bon sens des magistrates. Mais qui nous répondra que cette conscience et ce bon sens ne s’égarent pas?» (Voltaire, 1818, p. 229). These words allow us to understand how conscious the thinkers of the time were of the fact that the moment the judges of fact were completely freed from the system of legal proof, this freedom could degenerate into a “despotic arbitrariness”¹⁰².

It will come as no surprise, then, that an Enlightenment current advocated less radical reforms to the law of evidence. In this perspective, for example, the idea of creating a mixed system of legal proof and free evaluation of evidence emerged. One of the first proponents of such an idea was Gaetano Filangieri, who proposed the creation of an evidentiary model in which the judge would not be able to convict without the existence of a certain amount of evidence prescribed by law, while he would still have to acquit if the evidence did not produce in his mind a moral certainty of the suspect’s guilt (Filangieri, 1806, p. 149 ff.)¹⁰³. Such a mixed system was driven by a clear logic: it aimed to exploit the strengths of both the Roman-canon and the Enlightenment systems of evidence.

In this respect, it should be recalled that the idea of building systems based entirely on a synthesis of legal proof and moral certainty was indeed successful in certain areas of continental Europe, especially in the first half of the 19th century. Reference is made to certain German-speaking states, such as the Austrian Empire, where this approach—also known as the theory of negative legal proof (“*negative Beweisregeln*”)—has found its way into important legal texts¹⁰⁴. Despite its compromising nature, the experience with these completely mixed systems of evidence did not last long and was finally abandoned in the middle of the 19th century, or at least in the following decades (Nobili, 1974, p. 197 ff.)¹⁰⁵.

In actual fact, voices in favour of such a mixed model of evidence were heard in the French Constituent Assembly when it debated the *Décret concernant la police de*

¹⁰² See Barbacovi (1820, p. 44).

¹⁰³ See also Pagano (1824, p. 79 ff.). In the German literature see Feuerbach (1813, p. 132 ff.) and Mittermaier (1834, p. 83).

¹⁰⁴ See Ambos (2023, p. 174 f.); Daniele (2009, p. 121-123); Glaser (1883, p. 10 ff.); Mittermaier (1859, p. 94 f.); Nobili (1974, p. 189 ff.).

¹⁰⁵ However, for rules of this kind, which are still inherited from the German system, see Ambos (2023, p. 180). For Italy, see Daniele (2009, p. 123 ff.).

sûreté, la justice criminelle et l'établissement des jurées; i.e. the law that, by replacing the old *Ordonnance criminelle* of 1670, marked a historical turning point in continental criminal procedure and created a «unique meeting point between Anglo-Saxon procedure and the ideals of the French Revolution» (Nobili, 1974, p. 151). During the drafting of this law, one of the main points of contention between the progressive ideas of the Enlightenment and the opposition of a reactionary section of the Assembly was precisely whether to retain the classical rules of legal proof or to replace them with the standard of the jury's "intimate conviction" (*intime conviction*)¹⁰⁶, a concept which, in the French language of the time, was essentially synonymous with the idea of "moral certainty"¹⁰⁷. In this context, the theory of "negative legal proof" was supported by one of the main faces of the Revolution: Maximilienne de Robespierre. In fact, despite his distrust of the traditional legal system, he was convinced that it was necessary to «réunir et la confiance qui est due aux preuves légales et celle que mérite la conviction intime des juges», so that the accused could not be convicted on the basis of the former «si elles sont contraires à la connaissance et à la conviction intime des juges» (Robespierre, speech of the 4.1.1791, in Mavidal and Laurent (eds.), 1885, p. 11).

In France, however, this proposal was rejected in the very session in which it was formulated by the rapporteur of the bill: Adrien Duport¹⁰⁸. The decisive factor was the idea that the system of negative legal proof would have been incompatible with the institution of the jury, which, as we have seen, was a pillar of the Enlightenment ideal of the criminal process. Indeed, according to Duport, even such a model of fact-finding would have produced «de très mauvais juges au lieu d'excellents jurés» (Duport, speech of the 4.1.1791, in Mavidal and Laurent (eds.) (1885, p. 12), thus preventing the system from functioning properly from the outset.

5.2. The rise of the two-faced rule of *intime conviction*

The critics of the traditional Roman canon were ultimately successful. The *Décret* of 16-29 September 1791 abolished indeed the old system of legal proof and replaced it with the formula of the "intimate conviction"¹⁰⁹ of the jury. This expression was, in fact, enshrined both in the oath formula of the new law («vous jurez [...] de

¹⁰⁶ For a detailed study see Padoa-Schioppa (1994, p. 102 ff.).

¹⁰⁷ See, for example, Duport, speech of the 4.1.1791, in Mavidal and Laurent (eds.) (1885, p. 12); Robespierre, speech of the 2.2.1791 (*ibidem*, p. 718); Thouret, speech of the 11.1.1791 (*ibidem*, p. 132); Tronchet, speech of the 5.1.1791 (*ibidem*, p. 34), who used this or similar expressions, including that of "conviction morale" or "preuve morale". On the other hand, even in later literature, the concepts of *intime conviction* and *moral certainty* were considered by many to be synonymous, see, for example, Brugnoli (1846, p. 24); Hélie (1867, p. 14); Id. (1870, p. LXX); Lucchini (1895, p. 167). For contemporary literature, see Guérin (2015, p. 610 ff.) and Richard (2017, p. 20).

¹⁰⁸ See Duport, speech of the 4.1.1791, in Mavidal and Laurent (eds.) (1885, p. 11-13). But in defence of the old system of legal proof, see Goupil de Préfelin, speech of the 5.1.1791, (*ibidem*, p. 25).

¹⁰⁹ For a general discussion of this point, see Nobili (1974, p. 151-155).

vous décider d'après les charges et les moyens de défense et suivant votre conscience et votre intime conviction)» and in the explanatory instructions on their duties issued in the following weeks¹¹⁰. In this way, the centuries-old division between full proof and half proof was abolished in the heart of Europe, where the reformist ideas of the most progressive current of the Enlightenment were effectively implemented. From here, these ideas were transplanted, albeit with some adaptations and differences, into the mixed model of Napoleon's *Code d'instruction criminelle* of 1808 and have spread throughout the European continent to this day (Ambos, 2023, p. 171 s.).

Having clarified this, we can now look more closely at the twofold function of the formula *intime conviction* for the French Assembly of 1791¹¹¹.

From a first perspective, it constituted the standard of persuasion, indicating the degree of conviction that the jury had to reach at the end of the criminal trial in order to reach a guilty verdict¹¹². In this respect, the instruction explaining the duties of the jury, linked to the *Decree* of 1791, was particularly explicit, asking whether the impression «*ont fait sur le raisonnement les preuves apportées contre l'accusé et les moyens de la défense*» was such as to determine «*une intime conviction*» of the guilt of the accused. It should be noted, however, that the choice of the expression “*intime conviction*” rather than the more classical “moral certainty” has given rise to problems of no less importance, which go far beyond creating an apparent terminological distance between this parameter and those used in Anglo-Saxon countries, overshadowing their common origin. We are referring to the fact that, while the first term, especially thanks to its secular theological-Christian background, assumed a more easily definable scope, the second, especially when read in isolation, ended up being affected by a greater dose of vagueness. This is because, unlike other standards of persuasion, the formula of “intimate conviction” does not contain any words—such as the concept of certainty—that would clarify how firm the belief must be in order to be able to consider the accusatory hypothesis as sufficiently proven. On the contrary, the adjective “*intime*” merely draws attention to the internal forum of the individual. In so far as it is relevant here, these considerations are useful for understanding how, among the different standards of persuasion examined here, the *intime conviction* is the one that is affected by the greatest dose of linguistic vagueness, precisely because it does not contain any terms aimed at directly specifying the degree of solidity that the belief must assume in order to be considered proven. In view of this, it is not

¹¹⁰ See the *Instruction sur la procédure criminelle, en annexe de la séance du 29 septembre 1791*, published in Mavidal and Laurent (eds.) (1885, t. XXXI, p. 642 ff.), which states: «*la loi ne [...] leur prescrit point de règles auxquelles ils doivent attacher particulièrement la plénitude et la suffisance d'une preuve; elle leur demande de s'interroger eux-mêmes dans le silence et le recueillement, et de chercher, dans la sincérité de leur conscience, quelle impression ont faite sur leur raison les preuves [...]. La loi [...] ne leur fait que cette seule question [...]: avez-vous une intime conviction?*».

¹¹¹ On this point, see Delmas-Marty (1996, p. 59).

¹¹² In this sense, see Ferrer Beltrán (2021b, p. 19); Esnard *et al.* (2015, p. 133); Damaška (2022, p. 197 f.); Pradel (1985, p. 579); Thaman (2016, p. 82).

surprising that this rule is today strongly criticised by the advocates of the “objective” theory of the standards of proof¹¹³.

A second function assumed by the *intime conviction* in its original context was that of a general criterion for the evaluation of evidence according to the common sense of the jury, conceived in contrast to the classical system of legal proof¹¹⁴. From this point of view, the concept of *intime conviction* symbolised the freedom of decision-makers from the old-fashioned system of legal proof, which determined the value of evidence *a priori*. The above-mentioned instruction of 29 September 1791, which states that the law «ne [...] prescrit point des règles auxquelles ils doivent attacher particulièrement la plénitude et la suffisance d'une preuve», is also valuable for understanding this. Nevertheless, it is useful to note how, in the legal tradition of some countries, autonomous expressions have developed over time to refer to this second function of *intime conviction*. This is the case, for example, with Italian and German expressions such as “*libero convincimento*”, “*freie Beweiswürdigung*” or “*freie Überzeugung*”¹¹⁵.

While this is true, it must be pointed out that the rule of intimate conviction has given rise to significant exegetical problems, partly due to its semantic polyvalence. In fact, the formula in question has led both to restrictive interpretations (for example, it has been read only as a criterion for evaluating evidence and not as a standard of proof¹¹⁶) and to far much more dangerous expansive exegesis. This is the case, for example, in the Italian system, where the principle of *libero convincimento* has long been used by the judiciary as a caveat to overcome even certain limits set at the stage of admitting evidence, not at the stage of assessing it¹¹⁷. Obviously, in hypotheses such as these, the principle of free evaluation of evidence has been transformed from a rule created solely to close the door to the system of legal proof, into a formula that seeks to free the decision-makers from any kind of regulatory limitation imposed in order to circumscribe their discretionary power with regard to the establishment of the facts.

Given these problems, it is easy to understand why even one of the greatest European jurists of the nineteenth century, Carl Joseph Anton Mittermaier, considered the concept of intimate conviction to be «imprecise and confusing» (Mittermaier, 1851, p. 480). The reason for this opinion was the fact that this criterion, partly because of its explicit reference to the internal forum of the individual, ended up being perceived by a school of thought as an “oracular” rule that left the fact-finders so much room for autonomy that they could decide on the basis of their intuitions

¹¹³ See, for instance, Ferrer Beltrán (2021b, p. 19) and Laudan (2005, p. 98 ff.).

¹¹⁴ See Ambos (2023, p. 168).

¹¹⁵ For a recent study from the German perspective, see Ambos (2023, p. 168 ff.). In Italy it is still essential to read Nobili (1974). More recently, see Carlizzi (2018) and Nobili (2003, p. 33 ff.).

¹¹⁶ See, for instance, Lucchini (1895, p. 167); Taruffo (2003, p. 666) and Ubertis (2013, p. 332).

¹¹⁷ See Nobili (1974, p. 270 ff.).

and feelings (Damaška, 2019, p. 197; Glaser, 1883, p. 18). As Massimo Nobili has pointed out, this reading of the criterion of intimate conviction is, on closer examination, anti-historical: in fact, the very terms used in the 1791 instruction show how the French Constituent Assembly, for its part, inextricably linked the jury's intimate conviction to the impressions left in the minds of the jurors by the evidence (Nobili, 1974, p. 97 f.), thus attributing to it, in any case, an objective anchorage (Delmas-Marty, 1996, p. 60).

However, this interpretation of the formula was soon opposed by an antithetical one, which is now dominant on the European continent. We are referring to the "rationalist" interpretation of *intime conviction* (i.e. "*conviction raisonnée*"), which took root in the progressive wave of reflux that the institution of the jury experienced in many European countries, including France, after the splendour of the late eighteenth and nineteenth centuries (Damaška, 2019, p. 121 ff.)¹¹⁸. And then, as was to be expected, when systems based on professional judges or mixed courts broke the original link between intimate conviction and the jury, the need gradually arose to structure a theory of evidence, free from legal proof, but nevertheless based on respect for the rules of logic, science and reason. At the procedural level, the main legal institution that has been devised to pursue such an outcome is that of the *de facto* grounds of judgment (Thaman, 2016, p. 94 ff.). The existence of the duty to give reasons for a judicial decision, which has become a general pillar of European criminal law, means that judges must always be able to explain, in a way that is intersubjectively comprehensible and verifiable, both inside and outside the trial, why a factual hypothesis was or was not put forward. This system, which favours the use of solid criteria of inference in accordance with the state of scientific knowledge at the time, and at the same time with the indications given by the jurisprudence of the higher courts, obviously constitutes an essential safeguard against irrational attitudes on the part of judges towards the evidence¹¹⁹.

Looking more closely, there are early examples in Italian Restoration law of how some nineteenth-century legal systems went a step further by creating a mixture of rules concerning the standard of proof and the duty to give reasons for a judgement. One such example is the Regulation of Criminal Procedure of 5 November 1831, dictated by Pope Gregory XVI. Although it was marked by considerable inquisitorial and authoritarian incrustations (Contigiani, 2007, p. 189 ff.), this text could only appear very modern in terms of decision-making criteria. In fact, it was not limited to the obligation to give reasons for judgments¹²⁰, but also stated in Article 442, that at the end of the trial the professional judges must decide «according to the deep conviction of their conscience and according to the impression made on their minds by the

¹¹⁸ For Germany, see Ambos (2013, p. 176). For Italy, this reading was soon developed thanks to the thought of Romagnosi (1874, p. 665 ff.). On this point, see Nobili (1974, p. 53).

¹¹⁹ See *Taxquet v. Belgium*, n° 926/05, ECtHR, 16 November 2010. On this judgment, see Roberts (2011, p. 213 ff.).

¹²⁰ See Arts. 450, 452 and 465 of the Regulation.

evidence [...], on the gathering of which [...] must depend that fullness of moral certainty which removes all reasonable hesitation from their minds»¹²¹. These rules were an attempt to reconcile one of the main innovations of the French model of criminal procedure, namely the standard of *intime conviction*, with rationalist considerations aimed at preventing oracular drifts of this criterion and, finally, with the desire not to abandon the traditional “Christian” rules of “moral certainty” and the BARD. In short, it was an embryonic attempt to regulate, by means of procedural safeguards different from the classical rules of legal proof, norms that appeared subjective at the linguistic level.

It should be noted that in the history of Italian criminal procedure, more than a century and a half had to pass before a rule comparable to Article 442 of the Gregorian Regulation was codified again. In fact, after the unification of the country, the Italian codes of criminal procedure were for a long time strongly inspired by the French model of the *Code d'instruction criminelle*, which, if it favoured a definitive normative crystallisation of the standard of *intime conviction*¹²², on the other hand led to the disappearance of other references to moral certainty or reasonable doubt from the legal text for a long time. Nevertheless, it must be acknowledged how the formula of moral certainty has managed to survive for a long time, both in academic works¹²³ and in jurisprudence¹²⁴, thanks also to the historical link with French intimate conviction, gradually losing its theological-Christian background. After the Second World War, however, references to this criterion gradually diminished, although they never disappeared completely¹²⁵. In any case, both the notion of moral certainty and that of *intimo convincimento* have been undermined in the last half of the twentieth century by another “fossil” from the ancient past: the standard of beyond any reasonable doubt, which has been increasingly used both in theory¹²⁶ and in practice¹²⁷. This was the beginning of a long journey¹²⁸ that led to the codification of the BARD as a standard of proof for criminal convictions by the Italian Parliament in 2006¹²⁹. However, as we shall see below, this new codification cannot come

¹²¹ On this provision, see Ala (1839, p. 254 f.) and Giuliani (1840, p. 541 f.).

¹²² For the codes of 1865 and 1913, see respectively arts. 487/498 and 440 of the Code of Criminal Procedure, which were modelled on the old French system of 1791. Lastly, the *intime conviction* formula has remained unchanged and is still referred to in Article 30 of law 287 of 30 April 1951 on the *corte d'assise* (a mixed court composed of both ordinary citizens and professional judges).

¹²³ See Lucchini (1895, p. 167) and Saraceno (1940, p. 9).

¹²⁴ See Court of Cassation, judgment of 6 October 1883, n. 1368, p. 978.

¹²⁵ For a more recent use, see Court of Cassation, judgment of 26 May 2016, n. 42056, in *DeJure*.

¹²⁶ See Carnelutti (1963, p. 33).

¹²⁷ See Court of Cassation, judgment of 22 October 1984, Mattia and Court of Cassation, judgment of 24 February 1981, Pressi, p. 697 s.

¹²⁸ Two milestones along this path were the book by Stella (2003), and the judgment of the Joint Chambers of the Court of Cassation, of 10 July 2002, Franzese, p. 1133 ff.

¹²⁹ As a result, Article 533(1) of the Italian Code of criminal procedure now states: «the court shall deliver a judgment of conviction if the accused is proven to be guilty of the alleged offence beyond a reasonable doubt». On this reform, see Catalano (2016); Dalia (2018) and Pierro (2011).

as a surprise, since it is set in a historical context in which the BARD has gradually become a globally recognised standard of proof. In order to understand how this came about, however, it is appropriate to conclude our diachronic analysis by examining the evolution of the rule in question in common law countries.

6. THE US EXPERIENCE AND THE ENGLISH DIVERGENCE

Without the need to overturn a centuries-old system of legal proof applied by professional judges, or to compete with the formula of *intime conviction*, moral certainty and BARD have had less difficulty taking root in the Anglo-American model of trial by jury than on the European continent. In particular, a series of nineteenth-century treatises on the law of evidence, which had the merit of “secularising” these concepts, ensured their widespread dissemination (Shapiro, 1991, p. 25 ff.). A prominent role in this respect must be given, for example, to an influential work, both in England and in the United States: Thomas Starkie’s *Practical Treatise on the Law of Evidence* (1826). In this volume, the author took up the traditional distinction between “metaphysical certainty” and “moral certainty” and concluded that, in the criminal sphere, evidence should provide the trier of fact with «moral certainty to the exclusion of all reasonable doubt» (p. 514). But Starkie did not stop there, he also provided a definition of the standard of proof for a criminal conviction similar to that of Beccaria. For Starkie, a jury could not convict «unless [...] they were so convinced by the evidence that they would dare to act on that conviction in matters of the highest concern and importance to their own interest» (p. 514). In the United States, Simon Greenleaf (1842) agreed, saying that «the circumstances must be sufficient to satisfy the mind and conscience of the common man, and so convince him that he would dare to act on that conviction in matters of the highest importance and importance to his own interests» (p. 4 f.). What we have just seen shows how the ideas of the Enlightenment were also disseminated among Anglo-American lawyers. It was through the mediation of works such as those of Starkie, Greenleaf and others that these ideas were finally crystallised¹³⁰. To prove this point, suffice it to say that in both the United States and England, some of the explanations of the standard of proof for a criminal conviction still follow the tradition of Descartes and Beccaria, describing it as the state of mind in which a person «would act in his most important affairs»¹³¹.

¹³⁰ See, for example, Best (1849, p. 100), who, in order to demonstrate the universal recognition of the rule of moral certainty and beyond any reasonable doubt, made numerous references both to common law cases and to classical continental authors and sources, such as Art. 342 of the *French Code d’Instruction Criminelle*, which codified the rule of *intime conviction*. This shows how these concepts were perceived as equivalent.

¹³¹ For England, see *Mohammad* [2022] EWCA Crim 380 and *J.L.* [2017] EWCA Crim 621. For the US, see Supreme Court of Minnesota, *State v. Mitchell*, 16 April 1998, 577 N.W.2d 481, 485 (Minn. 1998) and Court of Appeals of Georgia, 23 June 1994, *Roura v. The State*, 214 Ga. App. 43

It was also thanks to this dogmatic background that the standards of reasonable doubt and moral certainty were established in American jurisprudence as early as the mid-nineteenth century (Laudan, 2006, p. 33). To facilitate their use, some judges began to develop articulate definitions of these terms, which gradually became enshrined in official instructions read to juries before deliberations. One of the most famous explanations of the BARD to emerge from this period is found in the 1850 Massachusetts Supreme Court case of *Commonwealth v. Webster*. In this case, Chief Justice Shaw, after admitting that the language of the rule was difficult to understand, sought to clarify its scope, stating that «reasonable doubt is not a mere possible doubt; because everything relating to human affairs, and depending on moral evidence, is open to some possible or imaginary doubt. It is that state of the case which, after the entire comparison and consideration of all the evidence, leaves the minds of the jurors in that condition that they cannot say they feel an abiding conviction, to a moral certainty, of the truth of the charge»¹³².

Insofar as it is relevant here, it is useful to note how this definition, which was so successful in some US jurisdictions that it is still used today, albeit with some adjustments¹³³, inverted the relationship between moral certainty and reasonable doubt. Whereas in the original theological and philosophical context, and still in Starkie's work, the state of mind to be attained was that of moral certainty, explained by the term reasonable doubt, in Shaw's instruction the situation was reversed: the standard had now become the BARD, which in turn was made explicit only by reference to moral certainty. This was a change, initially linguistic rather than conceptual, which was never subsequently challenged. In time, the BARD formula, which was preferred to the original lexicon also for terminological reasons - since it did not contain direct references to "morality" but to human "reasonableness" - was used to indicate the level to be reached for a criminal conviction.

In this respect, it is useful to recall that *Commonwealth v. Webster* is of particular importance, also because it established a link between the presumption of innocence and the BARD standard that has remained unchanged over time. According to Justice Shaw, precisely because everyone must be presumed innocent until proven guilty, it is not enough «to establish a probability, however strong [...] that the fact charged is more probable than not, but the evidence must establish the truth of the fact to a reasonable and moral certainty». At the end of the 19th century, this view was even accepted by the Supreme Court of the United States in the famous case of *Coffin v. United States*: at the end of a long historical excursus, full of cross-references to both the continental and common law traditions, this judgment consecrated the

(1994). In contrast, such instructions have been discouraged in New Zealand (*Wanhalla* [2006] NZCA 229) and Canada (*Lifchus* [1997] 3 SCR 320).

¹³² *Commonwealth v. John W. Webster*, 59 Mass. 295, 320 (1850). On this point, see Wharton (1880, p. 2).

¹³³ See Supreme Court of Massachusetts, *Commonwealth v. G. Russell*, 3 November 2014, 470 Mass. 464. For a critical analysis, see Welch (2013, p. 31 ff.).

link between the standard of reasonable doubt and the presumption of innocence, establishing that the former was legally derived from the latter¹³⁴. In this way, the BARD was able to anchor itself in a principle that «lies at the foundation of the administration of [...] criminal law»¹³⁵.

It is worth noting, however, that a similar link between the reasonable doubt standard and the presumption of innocence was also developed for a time in England, both doctrinally and jurisprudentially. In James Fitzjames Stephen's *History of the Criminal Law of England* (1883), for example, we read that the BARD rule is derived precisely from the presumption of innocence. On this point, the author - while openly admitting that «the word “reasonable” is indefinite» - did not fail to give the rule a specific function and meaning to the standard. Indeed, he stated, on the one hand, that «its real meaning [...] is that it is an emphatic caution against haste in coming to a conclusion adverse to a prisoner» and, on the other hand, that it could only be satisfied, justifying a man's conviction, if «every supposition not in itself improbable which is consistent with his innocence ought to be negated» (p. 438).

At a jurisprudential level, the most prominent English case that highlights the link between the standard of proof derived from the presumption of innocence and the BARD is *Woolmington v DPP*. In this landmark case, Justice Viscount Sankey uttered the famous words: in «English criminal law one always notes a golden thread: the prosecution has a duty to prove guilt [...]. If, at the end and as a whole of the case, there is reasonable doubt, created by the evidence given by the prosecution or the accused, [...] the prosecution has not proved guilt and the prisoner is entitled to acquittal»¹³⁶.

As regards the United States, it should be noted that the link between the reasonable doubt standard and fundamental principles of criminal justice has been further strengthened over time. In this respect, the famous *In re Winship* case is relevant, in which the Supreme Court held that the Due Process Clause, which can be derived from Amendments V and XIV of the Constitution, «protect[s] the accused against conviction except upon proof beyond a reasonable doubt of every fact necessary to constitute the crime with which he is charged»¹³⁷. In doing so, the Supreme Court established a direct link between the BARD and the US Constitution, elevating the standard to the top of the hierarchy of sources of US law¹³⁸.

In the decades that followed, it was once again the US Supreme Court that favoured a further step forward in the evolution of the standard for criminal conviction, helping to set some limits on the use of the ancient concept of “moral certainty”, the terminology of which has come to be perceived by many as archaic and, in some cases, even misleading (Laudan, 2006, p. 34 f.). In this respect, the 1990 *Cage*

¹³⁴ *Coffin v. United States*, 156 U.S. 432 (1895).

¹³⁵ *Coffin v. United States*, 156 U.S. 453 (1895).

¹³⁶ *Woolmington v D.P.P.* [1935] AC 462.

¹³⁷ See *In re Winship*, 397 U.S. 358 (1970).

¹³⁸ On this point, see Dripps (1987, p. 1665 ff.).

v Louisiana judgment is particularly relevant, which ruled against the use of the term “moral certainty” when combined with terms such as “grave uncertainty” or “actual and substantial doubt”¹³⁹. And while it is true that in the subsequent case of *Victor v. Nebraska*, the Supreme Court nevertheless confirmed that the use of instructions using the concept of moral certainty was not unconstitutional, it is also true that in the same judgment it recognised that «“moral certainty” standing alone, might not be recognised by modern jurors as a synonym for proof beyond a reasonable doubt»¹⁴⁰. These critiques provide a clear demonstration of how, in the American system, while moral certainty and BARD were originally two formulas used to refer to a single standard of decision-making, over time the latter concept gradually separated from its original core and took on an autonomous scope and meaning, not least because it was seen as more easily adaptable to contemporary society.

This is evidenced by the fact that, partly as a result of the Supreme Court’s decision not to endorse a single definition of reasonable doubt¹⁴¹, there has been a proliferation of other official definitions of reasonable doubt, many of which no longer refer to moral certainty, which, while not entirely abandoned¹⁴², has suffered a significant setback. In fact, depending on the state or federal district, judges give juries even very heterogeneous instructions on BARD: it is now conceived as a doubt for which a reason can be given; now as a high probability; now as a firm conviction; now as an achievable standard in the absence of alternative hypotheses about the subject’s guilt; now as a state of mind related to important decisions in one’s life; now as a state of mind that would cause a prudent person to hesitate to act, and so on (Laudan, 2006, p. 36-51; Shapiro and Muth, p. 57-59).

It will not go unnoticed that the choice of one or the other of these definitions is not without consequences, since, as even the US Supreme Court has recognised, the way in which a standard is explained can make it more or less demanding in practice¹⁴³. And it is precisely for fear of doing more harm than good by providing inadequate guidance on this point¹⁴⁴, that some American jurisdictions have taken the extreme step of banning the definition of BARD at the official level¹⁴⁵. It is worth recalling that this «sceptical point of view» (Tillers and Gottfried, 2006, p. 149) is linked to the thought of one of the most important American scholars of evidence:

¹³⁹ *Cage v. Louisiana*, 498 U.S. 39 (1990).

¹⁴⁰ *Victor v. Nebraska*, 511 U.S. 14 (1994). On this judgement, see Fortunato (1996, p. 365 ff.).

¹⁴¹ Indeed, in *Victor v. Nebraska*, 511 U.S. 5 (1994), the Supreme Court held that the Constitution neither prohibits trial courts from defining reasonable doubt nor requires them to do so as a matter of course.

¹⁴² But the formula still has its supporters: see Shapiro and Muth (2021, p. 69).

¹⁴³ See *Sullivan v. Louisiana*, 508 U.S. 275 (1993) and *Cage v. Louisiana*, 498 U.S. 39 (1990).

¹⁴⁴ It is worth noting that the US Supreme Court itself has stated that attempts to define reasonable doubt generally fail to make its meaning clearer to the jury: see *Holland v. United States*, 348 U.S. 140 (1954), and *Miles v. United States*, 103 U.S. 312 (1880).

¹⁴⁵ In this sense, see, for example, the Illinois Supreme Court in its decision of 18 June 2015, *People v. Downs*, 2015 IL 117934, 32.

John Henry Wigmore, who, starting from the assumption that «the truth is that no one has yet invented or discovered a mode of measurement for the intensity of human belief», concluded that «there can be yet no successful method communicating intelligibly to a jury a sound method of self-analysis of one's belief» (Wigmore, 1940, p. 325)¹⁴⁶. The problem is that Wigmore's criticism of attempts to verbalise reasonable doubt clashes with the fact that it is difficult to treat the BARD standard as a "self-evident" concept¹⁴⁷, because, as empirical studies have also shown, jurors end up attributing very different meanings to it (Smith, 2022, p. 294 ff.; Solan, 1999, p. 119 ff.).

Having reached this point, it is worth pointing out that it is precisely in order to remedy this set of problems that the legal system in England and Wales has, since the late 1940s, attempted to make an important departure in this respect (Keane and McKeown, 2019, p. 512 ff.; Roberts and Zuckerman, 2022, p. 275 ff.). It is an allusion to the fact that, although the reasonable doubt standard had been both ratified by the House of Lords on several occasions¹⁴⁸ and used in English criminal legislation¹⁴⁹, a movement developed to abolish it and replace it with a new phrase that was considered easier to apply and understand: the rule that the jury must be "sure" of the defendant's guilt. Chief Justice of the House of Lords Goddard played a crucial role in this turn of events, stating on a number of occasions that he preferred the use of the term "sure" to the old phrase "beyond any reasonable doubt"¹⁵⁰. And while it is true that this proposal failed to eliminate BARD from legal practice (McKeown, 2022), it must be acknowledged that since this pronouncement the standard of "sure" has indeed been endorsed by the majority of case law¹⁵¹ and by soft law. Indeed, since the late 1980s, the Judicial Studies Board (now the Judicial College) has also supported this view (Roberts and Zuckerman, 2022, pp. 275 ff.). It is no coincidence, therefore, that the latest edition of the *Crown Court Compendium* still refers to the "sure" standard rather than the BARD¹⁵².

However, it should be noted that the same Guide states that if a lawyer still refers to the criterion of reasonable doubt, «the jury should be told that this means the same as being sure» (Judicial College, 2023, § 5 n. 3). This makes it clear that the use of this term is not so much intended to change the level of the standard of proof with regard to the past (which would not have been permitted by a mere act of soft

¹⁴⁶ Already Bentham (1827, p. 74) wrote that «the language current among the body of people is, in this particular [context], most deplorably defective».

¹⁴⁷ But see the United States Court of Appeals for the Fourth Circuit in *Murphy v. Holland*, 776 F.2d470, 28 October 1985, which stated precisely that «the term reasonable doubt itself has a self-evident meaning comprehensible to the lay juror».

¹⁴⁸ See *Miller v. Minister of Pensions* [1947] 2 All ER 372.

¹⁴⁹ See PACE 1984, s. 76 (2).

¹⁵⁰ See *Kritz* [1950] K.B. 82, 88-90, as well as *Summers* [1952] 36 Cr. App. R. 14, 15.

¹⁵¹ See *Desir* [2022] EWCA Crim 1071, 198; *Mohammad* [2022] EWCA Crim 380; *Bogdanovic* [2020] EWCA Crim 1229; *Boaden* [2019] EWCA Crim 2284; *Miah* [2018] EWCA Crim 563.

¹⁵² See Judicial College (2023, § 5 ff.).

law such as the one in question), but rather to use a locution that is considered easier for juries to understand¹⁵³. The problem, however, is that the new English-language formulation of the standard of proof for a criminal conviction is no less vague than its predecessors (Keane and McKeown, 2019, p. 512 f.). This is confirmed not only by empirical studies showing that jurors find even this phrase difficult to understand¹⁵⁴, but also by the Court of Appeal of England and Wales itself, which finally admitted that «to define what is meant by “reasonable doubt” or what is meant by “being sure” requires an answer difficult to articulate and likely to confuse»¹⁵⁵. In the light of this, it is not surprising that some English scholars have suggested an even more radical route: abandoning both the reasonable doubt standard and the reference to the term “sure” and replacing them with entirely new instructions aimed at explaining to jurors in detail the standard of proof that must be achieved in order to convict a defendant (Keane and McKeown, 2019, p. 516 ff.). However, this theoretical position runs up against a significant practical problem: the fact that the view that judges should spend as few words as possible on the standard of proof, in order to avoid confusing the jury with poorly formulated definitions, is even more entrenched in England than in the US¹⁵⁶. This approach is further confirmed at an official level by the Crown Court Compendium, which states that «it is unwise to elaborate on the standard of proof» (Judicial College, 2023, § 5 n. 3). However, as Roberts and Zuckerman have rightly pointed out, it is impossible not to notice how this solution ultimately leads to an «abdication of judicial responsibility» (Roberts and Zuckerman, 2022, p. 279), making defendants pay for the system’s inability to provide a common minimum definition of the standard for a criminal conviction.

7. THE BARD’S AGE

The above-mentioned English attempt to abandon the reasonable doubt standard, at least linguistically, contrasts with an opposite trend that emerged globally in the second half of the twentieth century: it alludes to the fact that since then the BARD formula has met with increasing success in a growing number of legal systems.

At the European level, in order to have a quantitative dimension of the phenomenon, it is useful to refer to an empirical study published in the early 2000s by the Law Society of England and Wales, which shows that at that time fourteen

¹⁵³ See, on this point, Roberts and Zuckerman (2022, p. 278).

¹⁵⁴ See the interesting empirical studies made by Zander (2000) 1517 and Id. (2020, p. 18). See also Keane and McKeown (2019, p. 512).

¹⁵⁵ *Majid* [2009] EWCA Crim 2563.

¹⁵⁶ This view also goes back to Goddard: see *Hepworth and Fearnley* [1955] 2 QB 600, 603. See also *Yap Chuan Ching* (1976) 63 Cr App Rep 7, 11, where it was said that «if judges stopped trying to define that which is almost impossible to define there would be fewer appeals».

legal systems had already started to apply the beyond reasonable doubt rule in their criminal justice systems, while in most of the others standards linked to the *intime conviction* formula continued to be used¹⁵⁷. But BARD's success has become even more pronounced in the last twenty years: It should be noted that this phenomenon has occurred in two different ways. In a few countries, such as Belgium¹⁵⁸ and Italy, the legislator has codified this standard directly in the rules of criminal procedure. In most cases, however, this standard of proof has only been established by case law: this is the case, for example, in Spain¹⁵⁹, Portugal¹⁶⁰ and Germany¹⁶¹, where the courts have expressly used this formula in many decisions.

This new extension of the reasonable doubt standard in continental Europe has a twofold effect. On the one hand, it contributes to the fact that, in several countries, the formula of intimate conviction is losing one of its original functions, namely that of a standard of persuasion. On the other hand, this phenomenon helps to clarify the scope of the *in dubio pro reo* rule, which is derived from the presumption of innocence. As the German Federal Court of Justice points out, the introduction of the BARD makes it clear that only "reasonable doubt" and not every category of doubt can be used in favour of the accused¹⁶². And this is an advantage because it reflects the inherently fallible nature of evidential reasoning.

Having reached this point, it should be noted that such a European development of the BARD standard has been favoured by several factors, including not only the cultural influence that the American procedural model (of which, as we have seen, this standard is a pillar) has had on the continent since the end of the Second World War, but also the work of certain supranational institutions. Reference is made in particular to the position of the European Court of Human Rights, which in recent years has

¹⁵⁷ See Ede and Ford (2004, p. 37 f.).

¹⁵⁸ The reference here is to the law of 21 December 2009 reforming the court of assises, which removed from the *Code d'instruction criminelle* the references to the notion of *intime conviction* and replaced them with a double reference to the standard of *au-delà de tout doute raisonnable* (Articles 327(2) and 326(2)). It is interesting to note that even in France, some authors have stated that the *intime conviction* standard must be interpreted as equivalent to the BARD: see, for example, Pradel (2016, p. 414).

¹⁵⁹ See, for instance, Tribunal Constitucional, Sala Primera, 129/1988, of 16 June 1998 or Tribunal Supremo, Sala de lo Penal, n. 2254, of 21 September 2023. For a comprehensive analysis, see Alcácer Guirao (2021, p. 14 f.), who interestingly points out that the projects for reform of the Spanish criminal procedure system presented in recent years also make explicit reference to the BARD standard.

¹⁶⁰ See *Supremo Tribunal de Justiça*, proc. no. SJ200801100041985, of 10 January 2008. On this point, see Penim Pinheiro (2021, p. 15).

¹⁶¹ See *Bundesgerichtshof*, judgment of 16 February 2022, 2 StR 399/21; *Bundesgerichtshof*, judgment of 12 July 2017, 1 StR 535/16; *Bundesgerichtshof*, judgment of 12 January 2017, 1 StR 360/16; *Bundesgerichtshof*, judgment of 11 May 2017, 4 StR 554/16; *Bundesgerichtshof*, judgment of 21 November 2017, 1 StR 261/17. See Ambos (2023, p. 184).

¹⁶² See *Bundesgerichtshof*, judgment of 16 February 2022, 2 StR 399/21, which states that «nur solche Gründe, die zu vernünftigen Zweifeln in einer für den Schuldspruch relevanten Frage Anlass geben, einer Verurteilung entgegenstehen; nur dann ist „in dubio pro reo“ zu entscheiden».

clarified its traditional jurisprudence on Article 6(2) of the ECtHR (1950). Indeed, the European Court has no longer confined itself to deriving from this provision the “old rule” that the burden of proof is on the prosecution¹⁶³, but clarified this concept by explicitly referring to the BARD¹⁶⁴. It will therefore come as no surprise to learn that, precisely in the wake of the Strasbourg case-law, some countries have decided to abandon the old standard of *intime conviction* and replace it with the BARD¹⁶⁵.

Looking beyond Europe, it is even clearer that the BARD has now taken on a global dimension. Indeed, this standard has been codified in the criminal procedure codes of many national legal systems¹⁶⁶ and has been endorsed by many other supranational institutions, both regional (such as the Inter-American Court of Human Rights)¹⁶⁷ and universal. With regard to the latter, it should indeed be recalled that this standard was not only been proclaimed by the UN Human Rights Committee (2007) in its General Comment on Article 14 of the International Covenant on Civil and Political Rights¹⁶⁸, but is also crystallised in Article 66 of the Statute of the International Criminal Court (ICC, 2021). It is not difficult to see how this series of recognitions has entrenched the idea that this standard is a “natural” and “universal” human right. All this seems to prove that the present is the age of the BARD.

In contrast to this “triumphal march” is the difficulty that interpreters around the world have in attributing a precise meaning to the rule. It is no coincidence, for example, that the New Zealand Court of Appeal has explicitly referred to the «frustratingly indeterminate nature of the concept of proof beyond reasonable doubt» (Wanhalla [2006] NZCA 229), or that it has been suggested that the formula under consideration results in a fuzzy decision rule because «we are unable to provide precise and uncontroversial probabilities of any kind to associate with it» (Anderson *et al.*, 2005, p. 260). Or, again, that Wigmore described this standard as an «elusive and indefinable state of mind» (Wigmore, 1940, p. 317).

On closer examination, these opinions cannot come as a surprise, not only because of the complete break between the formula in question and the theological and moral context in which it was born, but also because of the linguistic vagueness of

¹⁶³ See *Barberà, Messegué and Jabardo v. Spain*, n° 10590/83, ECtHR, 7 December 1988, § 77.

¹⁶⁴ See *Kerimoğlu v. Turkey*, n° 29600/10, ECtHR, 6 December 2022, § 67 and *Ajdarić v. Croatia*, n° 20883/09, ECtHR, 13 December 2011, § 51, which qualified the circumstance that «the prosecution has to prove its case beyond any reasonable doubt» as a «basic requirement of criminal justice».

¹⁶⁵ This is the case in Belgium, in the context of *Taxquet v Belgium*, n° 926/05, ECtHR, 16 November 2010.

¹⁶⁶ An example of such a trend line is the People’s Republic of China, where a 2012 reform introduced BARD in Article 53 of the Criminal Procedure Code (following a further reform in 2018, this provision was moved to Article 55). This phenomenon has also been particularly pronounced in South America: see Article 340 of the Chilean Code of Criminal Procedure and Articles 359/402 of the Mexican Federal Code of Criminal Procedure.

¹⁶⁷ See *Manuela et al. v. El Salvador*, IACtHR, 2 November 2021, § 149.

¹⁶⁸ See Human Rights Committee (2007, p. 9, § 30).

the words that make it up. The term “doubt” is indeed ambiguous¹⁶⁹, while “reasonable” is even “contestable”, since «it is clear that it embodies a normative standard, but different users disagree about the detailed contents of that normative standard» (Waldron, 1994, p. 526 f.). And it is easy to see how the association of an ambiguous term with a contestable one can only make it more complex to assign a precise meaning to a proposition.

At this point it will become clear why, as mentioned at the beginning of this paper, a transversal movement of thought has developed in parallel with the worldwide diffusion of the BARD, with the aim of reducing its conceptual vagueness. Although this is true, it should be noted that the solutions proposed to rationalise the standard of proof for a criminal conviction are very heterogeneous, and only the two main approaches can be briefly discussed here¹⁷⁰.

A first line of thought, developed particularly since the 1970s in the context of the so-called New Evidence Scholarship¹⁷¹, attempts to solve the problem by using mathematical methods to calculate quantitatively the degree of “subjective probability” that a disputed fact is true, given the evidence gathered¹⁷². I am referring, in particular, to the proposal to use of the so-called “Bayes theorem”¹⁷³ to test whether the threshold for a criminal conviction has been reached¹⁷⁴. In short, proponents of this approach maintain that any evidential inference should be based on the application of Bayes’ rule, which is a useful tool for determining the rationality of changing one’s subjective belief in a hypothesis. Specifically, the fact-finder should use this theorem to convert his prior subjective probability (i.e., the degree of rational belief he has in a hypothesis before considering a particular piece of evidence) into the posterior probability (i.e., the probability of the same hypothesis after considering the evi-

¹⁶⁹ In fact, this term can be understood as: a) a subjective state of uncertainty, i.e. a belief or an opinion that is not sufficiently determined, or the hesitation to choose between an assertion and its negation; b) the insufficiency of reasons to believe in a certain fact - combined with an insufficiency of reasons to believe in an incompatible fact. On this subject, see Abbagnano (1998, p. 331).

¹⁷⁰ For further discussion see Pardo (2023, p. 431 ff.) and the special issue edited by Prakken, Bex and Mackor (2020, p. 1053 ff.).

¹⁷¹ On this point, see Lempert (2003, p. 91 ff.) and Twining (1991, p. 295 ff.).

¹⁷² On this notion, see Ferrer Beltrán (2007, p. 107 ff.), who reminds us that “subjective probability” is an epistemic notion of probability that measures the strength of our rational belief in a hypothesis in the light of some element of judgement.

¹⁷³ In short, Bayes’ theorem makes it possible to assess the impact of an evidential datum E on a fact-finder’s personal degree of belief in a given reconstructive hypothesis. In other words, the Bayesian method is used to understand whether and to what extent the presence of a given element strengthens or weakens a subject’s rational belief in the truth of a hypothesis. For a clear explanation of the application of the Bayesian method in the legal field, see: Taroni *et al.* (2023, p. 656 ff.).

¹⁷⁴ For a defence of the Bayesian approach to legal decision-making, see Eggleston, R. (2004, p. 169 ff.); Finkelstein and Fairley (1970, p. 489 ff.); Friedman, (1997, 276 ff.); Id. (2000, p. 873, ff.); Lempert (2003, 91 ff.); Garbolino (2014, 98 ff.); Kaye, D.H. (1999, p. 1 ff.); Nance (2016); Picinali (2022). For a practical application of Bayes’ theorem to check that the standard of proof in criminal cases is satisfied, see: Tribunal of Milan, judgment of 18 June 2015.

dence). At the end of this process of updating the degree of rational belief, the trier of fact arrives at the probability of the hypothesis given all the available evidence and should thus be able to consider whether the probability of the prosecution's hypothesis meets the level required by the standard for a criminal conviction¹⁷⁵.

The greatest strength of the Bayesian interpretation lies precisely in its claim to eliminate the vagueness inherent in the classical subjective language of standards of proof by relying not on legal rules, as in the continental tradition of legal rules of proof, but on the formal nature of mathematical calculation. Understood in this way, the standard of proof can indeed be made to coincide with a certain numerical “quantitative” threshold, which can be set at different levels depending on different normative choices (e.g. 99%)¹⁷⁶, beyond which the fact-finder is entitled to believe a hypothesis to be proven. This has an obvious theoretical advantage in terms of verifying *ex post* whether or not a certain threshold of evidence has been reached.

If this is true, it should be noted that this interpretation of the standard of proof for criminal convictions, even in its most sophisticated form based on Bayesian networks¹⁷⁷, is itself not free from several criticisms¹⁷⁸. Its main weakness is that the calculation on which it is based requires the establishment of a positive subjective probability of guilt *a priori*¹⁷⁹, a requirement which, apart from causing friction with the principle of the presumption of innocence¹⁸⁰, raises, above all, a serious problem of practical feasibility, since Bayes' theorem says nothing about how to assign one's degree of belief in the truth of the hypothesis of guilt before evaluating any evidence¹⁸¹. Nor, it should be pointed out, is this value usually identifiable in any other unambiguous way¹⁸², except perhaps in those rare cases in which the entire compendium of evidence itself to statistical analysis of the frequency of certain events that serve to circumscribe an initial “reference population”¹⁸³. And it is precisely because each proposed solution to the prior question in criminal trials raises serious theoretic-

¹⁷⁵ See on this point the clear analysis of Dahlman and Kolflaath (2021, p. 287 f.).

¹⁷⁶ For an interesting attempt to identify the appropriate percentage for identifying this threshold, see Cherubini (2024, p. 391 ff.), who, however, insists on the importance of using the Bayesian method to check the consistency of the evidentiary argument, but not to calculate the final probability of the defendant's guilt (p. 348).

¹⁷⁷ For a presentation of this method, see Taroni *et al.* (2014).

¹⁷⁸ In this regard, in addition to the essential work of Tribe (1971, p. 1329 ff.), see Allen and Pardo (2007, p. 107 ff.); Cohen (1977); Ferrer Beltrán (2021b, p. 80 ff.); Haack (2014, p. 56 ff.); Laudan, 2006, p. 100 ff.; Taruffo (1992, p. 168 ff.).

¹⁷⁹ For a detailed analysis of the problem of the prior, see Cherubini (2024, 346 ff.); Dahlman (2018, p. 15 ff.) and Dahlman and Kolflaath (2021, p. 287 ff.).

¹⁸⁰ See Cherubini (2024, p. 357 f.) and Dahlman and Kolflaath (2021, p. 289 f.).

¹⁸¹ See Makor *et al.* (2021, p. 456).

¹⁸² See, in particular, the convincing conclusions of Dahlman and Kolflaath (2021, p. 299).

¹⁸³ On this point, see Ferrer Beltrán (2021b, p. 82), who points out that in these cases the problem of «*adecuación al utilizar datos frecuenciales como modo de determinar probabilidades de proposiciones referidas hechos individuales*» remains.

cal and practical problems of its own, that the Bayesian methodology does not seem to be a good candidate for solving the problems of the BARD standard¹⁸⁴.

Given this, it is not difficult to understand why other approaches have been developed to rationalise the BARD, using qualitative rather than quantitative methods¹⁸⁵. One of the most promising theories proposed in this respect is based on a different and non-mathematical probabilistic reasoning: the so-called “Baconian” approach. This concept, which became popular in the legal field thanks in particular to the works of Laurence Jonathan Cohen (1977 and 1998), has as its fundamental characteristic that it focuses not on the degree of belief of the fact-finder, but on the degree of support that the evidence provides—objectively—for a hypothesis and, more specifically, on the idea of the “induction by elimination”, i.e. the progressive elimination of hypotheses that do not pass certain tests of evidential support¹⁸⁶. This is a fundamental aspect: «the weight of evidence in Baconian terms is related [...] both to the number of evidential tests a hypothesis survives and to the number of tests that might have been performed but were not» (Schum, 1994, p. 244). It is precisely because of this “objective” dimension that this theory is considered by many scholars to be one of the most suitable for rationalising evidential reasoning in a general way¹⁸⁷.

It should be recalled that the BARD, understood from an inductive probability perspective, is not to be interpreted as a numerically quantifiable (or cardinal) standard, but rather as a qualitative (or ordinal) threshold¹⁸⁸. This requires, on the one hand, a specific consideration of the extent to which the evidence fully covers the issues recognised as relevant and, on the other hand, the application of a progressive eliminative reasoning, discarding rival hypotheses in order to verify that the only

¹⁸⁴ It should also be pointed out that denying that Bayes’ theorem can be used as a useful tool for solving the problems of the criminal standard of proof does not mean that it cannot serve other important functions in the criminal process. One important example is to ensure that the assessment of the evidence is consistent with the premises of rational reasoning (Cherubini, 2024, p. 347), which can help legal fact-finders to avoid some common fallacies and biases in the evaluation of legal (and especially forensic) evidence (see, on this point, Dahlman, 2020, 1115 ff. and Makor, 2024, p. 101). In this respect, I agree with those who have argued that pluralistic approaches to the relationship between probability and evidence can be adopted, depending on the different objectives at hand: see, for example, Schum (1994); Kadane and Shum (1996, p. 152) and, more recently, Jellema (2023). See also Tillers (2011, pp. 167 ff.), who insists on the need to focus on the purpose for which formal methods are used in the process context.

¹⁸⁵ Alongside the Baconian method based on eliminative induction, which we will discuss in the text, there is a heterogeneous group of explanation-based/storytelling accounts of criminal evidence. For an analysis of these, see Makor et al. (2021, p. 431 ff.).

¹⁸⁶ On this point, see Anderson, Schum and Twining (2005, p. 257 ff.); Ferrer Beltrán (2021b, p. 88 ff.); Schum (1994, p. 243 ff.); Taruffo (1992, p. 199 ff.); Tuzet (2023, p. 127 ff.); Twining (2006, p. 125 ff.).

¹⁸⁷ See, for instance, Ferrer Beltrán (2021b, p. 88 ff.); Caprioli (2009, p. 66 ff.); Taruffo (2009a, p. 308); Tuzet (2003, p. 225 ff.).

¹⁸⁸ See Ferrer Beltrán (2007, p. 124).

one that can adequately explain the known elements is the hypothesis indicating the guilt of the accused. Indeed, it should be remembered that, according to a prevailing interpretation, the application of this theory to the BARD rule requires the fact-finder, before reaching a verdict of guilt, to eliminate every hypothesis that is consistent both with the evidence and with the innocence of the accused¹⁸⁹.

This theory, like other explanation-based methods of reasoning about evidence, does not clearly allow us to achieve a level of precision analogous to that of Bayesian cardinal proposals¹⁹⁰. But it has other virtues. Among these is the fact that it allows us to avoid the conceptual and operational problems that characterise the subjective concept of probability, including that of assigning probabilities *a priori*, since it is not necessary to establish such a value as the starting point of a calculation¹⁹¹. The most important advantage, however, is that the method of eliminative induction makes it possible to rank competing hypotheses according to a graded scheme based on intersubjectively verifiable criteria: in particular, the internal consistency of a hypothesis, its consistency with background knowledge, the greater or lesser support it receives from the evidence, and, above all, its ability both to explain existing data and to predict new data, integrating them according to a coherent scheme¹⁹². And this is obviously a great advantage for rationalising the judgement phase, since such criteria allow a preponderant focus not on the beliefs of the adjudicators, but on how complete the evidential coverage of competing hypotheses is¹⁹³.

Having reached this point, it is now time to consider whether strategies such as these are sufficient to address the long-standing problems of standards of proof for criminal convictions, or whether more complex approaches are needed.

¹⁸⁹ See, for example, Accatino (2011, p. 507 f.); Cohen (1977, p. 249 f.); Mazza (2012, p. 366 f.); Roberts and Zuckerman (2022, p. 281 f.); Taruffo (2009a, p. 308). It is worth noting, however, that even non-probabilistic proposals for clarifying the standard of proof, such as the so-called relative plausibility theory based on the method of inference to the best explanation, arrive at results that are not too far off. See, for example, Allen and Pardo (2019, p. 16) who state that BARD would be satisfied only if the prosecution's explanation is plausible given the evidence, while there is no plausible defence explanation.

¹⁹⁰ It must be admitted that even after such a reinterpretation, given the merely ordinal rather than cardinal nature of the method of eliminative induction, it may be difficult in practice to know when one is faced with an alternative hypothesis that is sufficiently well-founded to leave the question of acquittal open. For a proposed solution to this seminal problem, see Roberts and Zuckerman (2022, p. 282 f.), who state that the hypothesis of innocence can be discarded if it can be said to be "inert", i.e. if it cannot reasonably be assigned even minimal probative value.

¹⁹¹ See Ferrer Beltrán (2007, p. 124 ff.).

¹⁹² See Tuzet (2023, p. 134 ff.) for a detailed analysis of these parameters.

¹⁹³ Schum (2009, p. 220) points out that, according to Cohen, evidentiary completeness is the most important factor associated with the weight of evidence.

8. CONCLUSIONS: WHAT STANDARD OF PROOF FOR A CRIMINAL CONVICTION FOR THE FUTURE OF CRIMINAL JUSTICE?

This essay has shown how different approaches to structuring an efficient standard of proof for a criminal conviction have evolved over the centuries. The “objective” approach typical of the continental Roman-canon system, based on reliance on a set of legal proof rules, was contrasted in the late modern period with a “subjective” approach, oriented towards the valorisation of common human reason. And it was precisely in the transition from the first to the second perspective that three of the standards that are still most used were definitively imported into the field of law: moral certainty, the BARD, and *intime conviction*. As we have seen, they find an original common extra-legal core in the *certitudo moralis* of the theological-Christian ascent: a standard developed in the Middle Ages to resolve cases of conscience. From this point of view, it is fair to say that these rules are in fact living “fossils” of an ancient past (Whitman, 2008, p. 203). But it was not until the Age of Enlightenment that they became firmly established in the field of law. This is no coincidence, since it was a particularly favourable period for such criteria, which are strongly linked to a belief in the ability of ordinary people to make reasonable decisions on controversial issues.

The problem is that while such rules were initially easy to understand, partly due to their familiarity with the language of the time, over the centuries they have become increasingly difficult to read. And it is precisely to resolve this impasse that various solutions have been proposed. The most radical one is to abandon the classical standards of proof for a criminal conviction in favour of new ones, structured at the semantic level, so that they no longer depend on the degree of persuasion of the fact-finder, but on the level of epistemic confirmation that the evidence provides for the hypothesis of the accusation¹⁹⁴. In other words, according to this interpretation, the era of decision-making rules based on the common sense of the Enlightenment should definitively come to an end and a new era marked by the creation of entirely new rules should begin.

¹⁹⁴ Reference is made to the proposals of Ferrer Beltrán (2021b, p. 208 ff.) and Laudan (2006, p. 82 ff.). The former, for his part, provided a list of seven standards of proof that could be applied in different areas of law, of which the one relating to criminal matters should be worded as follows. «*Para considerar probada una hipótesis [...] deben dars conjuntamente las siguientes condiciones: a) La hipótesis debe ser capaz de explicar los datos disponibles, integrándolos de forma coherente, y las predicciones de nuevos datos [...] deben resultado confirmadas y aportadas como pruebas al proceso. b) Deben haberse refutado todas las demás hipótesis plausibles explicativas de los mismos datos que sean compatibles con la inocencia del acusado [...], excluidas las meras hipótesis ad hoc*». Laudan, on the other hand, proposed three alternative formulations of a criminal standard: a) if there is credible, inculpatory evidence or testimony that would be very difficult to explain if the defendant were innocent, and there is no credible, exculpatory evidence or testimony that would be very difficult to explain if the defendant were guilty, then convict. Otherwise, acquit. b) If the prosecutor’s story about the crime is plausible and you can conceive of no plausible story that leaves the defendant innocent, then convict. Otherwise, acquit. c) Figure out whether the fact established by the prosecution rule out every reasonable hypothesis you can think of that would leave the defendant innocent. If they do, convict; otherwise, acquit.

Unfortunately, it should be noted that this proposal also has serious limitations. In the first place, in its own time, it clashes with the inherent vagueness of verbal language in being able to identify a threshold of evidential sufficiency¹⁹⁵. Indeed, there is a danger—denounced by other scholars—that even the new standards will be no more determined on the linguistic level than the previous ones, creating only new theoretical and practical problems¹⁹⁶. What makes such an approach very difficult, however, is above all the success that the standard of reasonable doubt in particular is enjoying at the political and jurisprudential level in many legal systems; a success which, if it can lead one to predict a further retreat of the other standards, makes it prohibitive to think that it too can really be abandoned, at least in the short term, in favour of another formula. Finally, I would add to these considerations the fact that many of the most advanced proposals for new standards are very similar to the rational reinterpretations of the BARD standard actually used in some legal systems¹⁹⁷. This raises the question of whether it is really worth taking the difficult political step of abandoning traditional decision-making criteria in order to adopt others that should have approximately the same meaning.

This leads me to prefer a less theoretically ambitious but in the short term more realistic approach, which is to maintain the traditional standards of proof for a criminal conviction, especially in countries—such as Italy and the United States—where the BARD rule has been given constitutional status¹⁹⁸. At the same time, however, it is necessary to try to reduce the risk that the classical standard of proof ends up being understood as a mere rhetorical invitation to the triers of fact to be particularly cautious when they are called upon to decide whether or not an accused is guilty of a crime¹⁹⁹.

To this end, I find particularly useful the opinion of those who have pointed out that, in order to solve, at least in part, the problem of the linguistic vagueness of the

¹⁹⁵ Ferrer Beltrán (2021b, p. 203) admits that the problem of linguistic imprecision in standards cannot be completely eliminated, but only marginalised.

¹⁹⁶ See Allen (2013, p. 55); Garbolino (2014, p. 484 f.); Tuzet (2023, p. 270 f.); Uberris (2013 p. 332 f.).

¹⁹⁷ I am referring to the fact that both Ferrer and Laudan insist on the need to exclude alternative hypotheses in which the accused might be innocent, as in the Baconian reinterpretations of BARD.

¹⁹⁸ In any case, I want to stress that I think the BARD is a better formulated standard than the others that have evolved historically. There are three reasons for this. First, it does not refer linguistically to potentially absolute mental states (such as moral certainty or the English sure standard), and this fits well with the fallible nature of evidential reasoning. Second, it does not contain explicit references to the internal forum or to morality, but to human reason. Finally, by requiring the “exclusion” of reasonable doubt, it lends itself to interpretation from a perspective related to eliminatory induction, i.e. the elimination by progressive evidence of the possibility that there are elements capable of founding a hypothesis on the disputed facts for which the defendant may be considered innocent. For a recent defence of the standard, based on a pragmatist understanding, see Tuzet (2024, p. 398 ff.).

¹⁹⁹ On this point, see Ho (2008, p. 185 ff.). In any case, this function alone is valuable because it is able to neutralise the negative impression created by the mere fact of the accusation: see Robert and Zuckerman (2022, p. 285).

standards of proof for a criminal conviction, the most important thing is to clarify the reasoning procedure to be followed in order to understand whether or not the standard of proof has been met in a particular case²⁰⁰. While, in fact, a threshold of proof can only be defined in general terms at the linguistic level, the method to be followed in order to reach this threshold can be specified much more precisely²⁰¹. To sum up, I think it is important to focus on the fact that the threshold of evidential sufficiency set by the BARD rule can only be said to have been reached if a certain method of reasoning is actually followed in the judgement. And it is clear that the aim of this method must be to preserve the preference for false acquittals over false convictions, which, as we have seen, is one of the hard cores of the presumption of innocence²⁰².

²⁰⁰ See, for instance, Picinali (2015, p. 139 ff.); Robert and Zuckerman (2022, p. 280 ff.). In the Italian literature, see Conti (2020, p. 829 ff.); Iacoviello (2006, p. 3873 ff.) and Id. (2023, p. 434 f). Interestingly, even Ferrer's and Laudan's attempts to objectively rewrite the standards of proof are more attempts to set qualitative reasoning methods for adjudicators to follow than attempts to set abstract evidentiary thresholds to be met.

²⁰¹ I am aware that there are at least two possible drawbacks to the suggestion of interpreting standards of proof as methods of reasoning. The first is that it might encourage a dogmatic confusion between legal proof rules and standards of proof, a confusion that already exists at various levels, including at the level of the European Court of Human Rights (see *Gäfgen v. Germany*, n° 22978/05, ECtHR, 1 June 2010, § 92; on this point see Tuzet, 2021, p. 91). However, this risk should be mitigated by the fact that the standard of proof should in any case indicate a useful method of reasoning for the evaluation of all types of evidence, and not only a specific category (such as testimony, confession or circumstantial evidence). The second possible objection has to do with the vagueness that is also likely to plague this proposal: in fact, identifying a method of reasoning about evidence in the standard of proof still tells us nothing about which method exactly should be used. Well, as I will soon reiterate in the text, in this regard I believe that in criminal law the guiding light that must indicate the method to be applied is the presumption of innocence. I mean, in particular, that the method must be able to incorporate the preference for false acquittals over false convictions, which is a pillar of this principle. It is then up to other legal institutions—such as the *de facto* grounds of judgment in civil law jurisdictions or the instructions to juries in common law jurisdictions—to specify this method of reasoning. Finally, it is also up to other legal institutions—such as appeal procedures—to make the method more effective in practice.

²⁰² It is worth noting that, also because of the link between the standard of proof and the presumption of innocence, I do not agree with those who have proposed that the standard of proof for a criminal conviction should vary according to the nature of the case and, in particular, according to the difference in the disutility of false positives and false negatives that can be attributed to different classes of criminal behaviour (see, for example, Ferrer Beltrán (2021b, pp. 154 ff. and p. 227)). I object to this argument, which is motivated by a logic similar to that of the establishment of “extraordinary punishments” mentioned above (§ 3.2), because if the standard of proof is indeed inextricably linked to the *ratio* of false positives to false negatives that a given jurisdiction considers acceptable for the application of the presumption of innocence, then this standard must be the same wherever the presumption is applied. Otherwise, some defendants would enjoy less protection of the presumption of innocence than others, even though it is the most important ethical-legal principle of criminal law. There is another consideration which leads us to believe that standards lower than the BARD cannot be used to decide whether or not to convict for crimes punishable by less severe penalties (such as fines). In fact, the standard of proof protects not only against punishment in the strict sense, but also against the social stigma of a criminal conviction. In other words, it is not only the punishment that counts, but also everything that derives socially from the status of a convicted person (such as greater difficulty in finding a job or the loss of

In so far as it is relevant here, the path to the construction of this reasoning procedure consists of several interrelated steps, the first of which requires that BARD be purified of the intuitionistic and moral encrustations that refer to the context in which it was born, and that it be brought up to date by the use of concepts that belong to contemporary epistemology. I believe that one of the best candidates from this point of view is the above-mentioned Baconian theory of “eliminative induction”, understood as the constant testing of hypotheses on facts, checking their internal consistency, their coherence with basic knowledge, the greater or lesser support they receive from the evidence, and their ability both to explain existing data and to predict new ones, integrating them according to a coherent scheme²⁰³. The Italian practical experience with the BARD seems to confirm this. Although I cannot analyse here the diachronic development of this rule in Italy²⁰⁴, I would like to point out one thing: since its slow rediscovery in the second half of the twentieth century, the concept of “reasonable doubt” in this country has acquired a more objective interpretation than the classical one linked to the concept of moral certainty (Mazza, 2025, p. 10). Since the formal codification of the BARD in 2006, this approach has been further consolidated by the joint efforts of scholars²⁰⁵ and jurisprudence²⁰⁶ to provide a sufficiently precise method for distinguishing between “reasonable” and “unreasonable” doubts. In order to resolve this issue, the concept of reasonable doubt has in fact been inextricably linked to the passing of a series of tests relating to the completeness of the existing evidential framework and to the presence or absence of evidence capable of excluding any explanation consistent with the innocence of the accused. In fact, only “tangible” doubts, i.e. doubts about the evidence or its absence,

certain state benefits or subsidies). From the opposite perspective, I do not think either that it is possible to set higher standards than BARD in the case of particularly harsh sentences (see, on this point, Lillquist, 2005, p. 45 ff.): and this because, as we have already noted, imposing a standard close to absolute certainty would risk preventing a penal system from functioning properly, given the inherently fallible nature of any evidentiary system. Rather, to better protect the innocent in the face of particularly severe penalties, it is more efficient to rely on other procedural safeguards, by considering higher standards of protection in terms of rules of assessment or exclusion of evidence, as well as by requiring unanimity and collegiality in the decision, or by ensuring greater rights of appeal.

²⁰³ Since, as I said earlier, different ways of thinking can be useful in rationalising reasoning about evidence, I believe that Bayesian methods could also be useful in this review process, particularly to reduce the risk of the trier of fact falling into logical or probabilistic fallacies. The same is true for explanatory theories that focus on the concept of “plausibility” (for an interesting examination of the meanings of this concept, see Dahlgren, 2024, pp. 91 ff.), as they are useful for assessing the degree of confirmation of different possible hypotheses about facts. In conclusion, I believe that, in order to solve the problem of the vagueness and subjectivism of the standard of proof for a criminal conviction, it is useful to adopt a pluralistic method of reasoning based on several elements. The main method is that of eliminative induction, but it must and can be supported by bayesian and plausibilist theories.

²⁰⁴ For a recent summary, see Neri (2024, p. 49 ff.).

²⁰⁵ See, for example, Caprioli (2009, p. 66 ss.); Catalano (2016); Conti (2020, p. 829 ss.); Daniele (2009, p. 172); Iacoviello (2006, p. 3873 ff.); Incampo and Scalfati (2017); Mazza (2012, p. 366 ss.); Taruffo (2009a, p. 310 ss.).

²⁰⁶ For a summary of the case law, see Triggiani (2017, p. 325 ff.).

are considered “reasonable”²⁰⁷. On the contrary, doubts that are merely “sceptical” or “intangible” in the light of the compendium of evidence are not considered “reasonable”²⁰⁸. It is precisely in this sense that the prevailing interpretation of the standard indicates that the decision rule laid down in Article 533 of the Italian Code of Criminal Procedure can be said to have been reached only if the evidence allows the judge to exclude the recurrence of two categories of doubts: a) “internal”, i.e. those which reveal the contradictory nature or the explanatory incapacity of the hypothesis of guilt; b) “external”, i.e. based on the existence of an alternative hypothesis of innocence which is not merely a speculative possibility but is actually plausible²⁰⁹.

But this first theoretical step is not enough. In order to make the standard of proof work practically and effectively, it is essential to surround it with a set of other rules which, on the one hand, are capable of helping to identify the method of reasoning which is intended to ensure that the relevant evidential threshold is respected and, on the other hand, are capable of ensuring that it is respected in practice.

As we have seen, in civil law systems with professional judges, the main safeguard which makes it possible to pursue this objective is the existence of rules on the *de facto* grounds of judgements, especially if this is accompanied by a system of appeals which enables the parties to challenge the incorrect application of the standard of proof. In this case too, the Italian example is particularly useful: the emergence of the BARD exegesis examined above was in fact favoured by that articulated system of motivation of judicial decisions, formed by a mixture of the obligation of dialogical motivation imposed on the judge (art. 192 and 546 of the Code of criminal procedure) and a complex system of remedies, both factual (the appeal under articles 593 et seq. of the Code of criminal procedure) and legal (the appeal to the Court of Cassation under articles 606 et seq. of the Code of criminal procedure), which can be activated by the parties²¹⁰. Indeed, it is clear how the need to explain why the standard was or was not met, coupled with the possibility of *ex-post* censure for failing to meet the required threshold and/or wrong motivation on this point, simplifies efforts to progressively specify the scope of the BARD formula. All this shows that the relationship between standards of proof and the motivation of judgments is twofold. On the one hand, standards are essential to enable triers of fact to justify their decisions in a non-arbitrary way. On the other hand, the imposition of specific motivational burdens favours a path of clarification of the threshold to be reached.

²⁰⁷ See the judgment of the Court of Cassation of 19 June 2018, n. 48541, in *Ced. Cass.*, n. 274358-01.

²⁰⁸ An important example is the judgment of the Court of Cassation of 21 May 2008, Franzoni, in *Ced. Cass.*, n. 240763, in which it is stated that the BARD requires a conviction to be pronounced when the evidence obtained leaves out only remote eventualities whose concrete realisation in the case is not reflected in the evidence at trial, placing them outside the natural order of things and normal human rationality.

²⁰⁹ See, *ex multis*, the judgment of the Court of Cassation of 24 October 2011 n. 41110, in *Ced. Cass.*, n. 251507, and, more recently, of 2 October 2023 n. 39777, in *DeJure*.

²¹⁰ On this point, see Conti (2020, p. 830 ff.).

In this context, it is clear that the greatest difficulty in attributing a precise meaning to the BARD in common law systems is also explained by the fact that juries decide in secret with an unmotivated verdict (see Taruffo, 2009a, p. 307 and Thaman, 2016, p. 75). Indeed, in such a context, there is no fundamental means of external feedback on the jury's interpretation of the standard of proof²¹¹. And it is precisely for this reason that, as the European Court of Human Rights pointed out in the *Taxquet v. Belgium* case, it is particularly important in such systems to value at least the element of the professional judge's instructions to the jury²¹². Against this background, it is difficult not to agree with those who have argued that it is necessary to abandon the approach, which seems to be widespread in both the English and the American legal systems, of treating decision-making criteria as self-evident concepts to which as few words as possible should be devoted. On the contrary, in order to avoid perpetuating the long-reported situation where juries have chronic difficulty in understanding the conditions under which they can convict, it is important to adopt a model of instruction that (i) defines the standard to be achieved in a way that enables jurors to appreciate their function in protecting the innocent from the moral harm of wrongful conviction, and (ii) makes clear to jurors how they should apply this crucial criterion. For example, one valuable option is to state in the instruction that «the jury may convict only when all explanations of the evidence that are consistent with innocence have been dismissed as untenable» (Roberts and Zuckerman, 2022, p. 281)²¹³. It will not go unnoticed that such a definition, again based on the scheme of eliminatory induction, is very similar to the reinterpretation of reasonable doubt used in some civil law countries, such as Italy, which shows how the distance between continental and Anglo-American legal systems in this respect can be relativised, at least in part, by using the right instructions.

Third, it should not be forgotten that the efficiency of the standards of proof is also decisively influenced by other substantive and procedural rules. The first point to note is that criminal offences must be constructed in a way that is consistent with the presumption of innocence, which—in principle—places the burden of proof on the prosecution. And from this point of view, the forms of legal clauses that reverse the *onus probandi* and place it, at least in part, on the suspect or defendant are problematic²¹⁴. This is for the obvious reason that the effective ability of the standard of proof to protect the innocent from wrongful conviction is in danger of being under-

²¹¹ It is true, however, that those forms of appeal that allow convicted persons to challenge the misapplication of the proof beyond a reasonable doubt requirement, such as the sufficiency of the evidence appeal in the US, provide a minimum guarantee of compliance with the standard. On this point, see Saltzburg and Capra (2018, pp. 1702 f.), who discuss the seminal case of *Jackson v. Virginia*, 443 U.S. 307, 309, 323-24 (1979).

²¹² See *Taxquet v. Belgium*, n° 926/05, ECtHR, 16 November 2010, § 92.

²¹³ As we have already mentioned, Stephen (1883, p. 438) had already expressed himself in similar terms.

²¹⁴ On this point see Ferrer Beltrán (2021b, p. 166 f.); Hamer (2007, p. 142 ff.); Roberts (2014, p. 317 ff.); Stumer (2010, p. 98 ff.).

mined by the existence of such forms of reversal clauses, which are still widespread in contemporary criminal justice systems²¹⁵. It follows that, while the European Convention on Human Rights does not prohibit the construction of crimes by taking into account forms of presumptions of fact or law *in malam partem*²¹⁶, it is essential that such a form of reversal of the burden of proof is not only always relative and limited to what is strictly necessary, but also always consistent with the standard of proof deemed optimal by the particular system. More specifically, this means that in a system based on the BARD rule, in order to rebut such presumptions, it should in principle be considered sufficient for the defence to raise even a reasonable doubt on the point, since it seems incoherent and dangerous to impose a higher standard of proof.

At the procedural level, account must be taken of all those rules which, by making it possible to identify situations of insufficient and/or contradictory evidence, favour the emergence of a reasonable doubt as to the innocence of the accused. From this point of view, both the rules on cross-examination and the right of defence (which must also be understood as the right to defend oneself by producing evidence²¹⁷) play a key role and require criminal proceedings to be structured in such a way that the accused is exposed to constant attempts to falsify the accusation. The same is clearly true, however, of a number of rules that specifically concern the judicial decision phase: among the latter, the most prominent are those rules—increasingly marginalised in many legal systems for reasons of procedural economy—that confer the power to decide on collegial bodies, especially when the latter are required to reach a unanimous decision. It is not difficult to see how the need for a panel, rather than an individual to decide whether the threshold of sufficiency of the evidence has been reached is itself a safeguard that makes the achievement of a standard far more complex²¹⁸.

What has just been said makes it possible to understand a fact that is too often underestimated. If it is true that the presumption of innocence requires the establishment of a standard of proof for a criminal conviction that favours the acquittal of the guilty over the conviction of the innocent, it is equally undeniable that in reality it is many other rules (substantive or procedural) that contribute to making the standard of proof truly concrete and effective (Ferrer Beltrán, 2021b, p. 166 f.). For this reason, it can be assumed that the standard of proof, even if it is equivalent from a

²¹⁵ See *Falk v. Netherlands*, n° 66273/01, ECtHR, 19 October 2004, which states that presumptions of fact or law operate in every criminal justice system.

²¹⁶ See in this respect the landmark judgment of the *Salabiaku v. France*, n° 10519/83, ECtHR, 7 October 1988, § 28, which states that presumptions of fact or of law having the effect of reversing the burden of proof are not in principle prohibited by the Convention, as long as States remain within certain limits, taking into account the importance of the matter in question and respecting the rights of the defence. See also Recital 22 of EU Directive 2016/343 on the presumption of innocence.

²¹⁷ See Vassalli (1968, p. 12).

²¹⁸ It is no coincidence that in *Ramos v. Louisiana*, 590 U.S. (2020), the U.S. Supreme Court, overturning an earlier precedent, reaffirmed that the unanimity requirement for felony convictions derives directly from the Sixth Amendment to the U.S. Constitution.

linguistic point of view, is in reality capable of protecting the fundamental rights of the accused in a more or less effective way, depending on the criminal justice system in which it is inserted and even on the individual procedural module applied in practice. For example, it will be easy to understand how, in principle, it was much easier to satisfy the criterion of moral certainty or reasonable doubt in the systems of the *Ancient régime*, which, by relegating the defence or cross-examination to a marginal role, drastically reduced the probability of the emergence of alternative innocent hypotheses or logical holes in the prosecution's hypothesis, than it is today, where such safeguards are generally crystallised as constitutional principles. But similar considerations apply to those contemporary legal systems that, while relying on the BARD standard for decision-making, unduly restrict the fundamental procedural rights of individuals. All this shows that legal systems must always be aware that the moment they sacrifice essential safeguards in order to falsify the prosecution's case, they end up indirectly affecting the standard of proof by substantially altering the way it can concretely distribute the *ratio* between false positives (conviction of an innocent person) and false negatives (acquittal of a guilty person)²¹⁹. In conclusion, in order to make the reasoning behind the standard of proof practical and effective, the criminal justice system as a whole must be taken into account.

And it is precisely because of the inherent fallibility of even contemporary legal systems that it is essential to ensure that the standard of proof required for a conviction is respected, not only until a judgment has become final, but also in the context of the extraordinary appeals or collateral attacks that may be brought to challenge convictions that have already become *res iudicata*. In particular, reference is made to the need to move away from the approach that has long prevailed in various legal systems, which tends to allow a final conviction to be overturned only if there are elements that positively establish the innocence of the convicted person²²⁰. In order

²¹⁹ It is not difficult to see that some of the most problematic legal institutions from this point of view are those of negotiated justice (such as plea bargaining), which are themselves becoming increasingly widespread at the global level. Plea agreements are particularly dangerous because they are usually based on the logic of exchanging benefits for the defendant's renunciation of a more or less extensive set of fundamental procedural rights. It is no coincidence that scholars have documented the so-called "innocence problem" of negotiated justice, i.e. the risk that parties who enter into an agreement with the state are in fact innocent. See, for example, Duce (2024, p. 278 ff.). From this point of view, it can be said that the rise of negotiated justice is capable of significantly weakening the above-mentioned international expansion of the BARD standard; and this due to the fact that in a large part of the legal systems this standard is either completely abandoned or, in any case, severely weakened after a plea bargain has been reached. To partially remedy this inconsistency, a change of perspective seems appropriate, aimed at ensuring that a set of inalienable guarantees is always recognised, even when the defendant opts for a plea bargain procedure: for some suggestions in this regard, see Della Torre (2019, pp. 595 ff.).

²²⁰ This was the case in Italy, for example, when the 1930 Code of Criminal Procedure was in force, Article 555 of which provided that, in order for an application for revision to be granted, new evidence had to be produced to show that acquittal was clearly necessary. With the adoption of the 1988 Code, this situation changed and the jurisprudence now recognises that for the convicted person

to create an effective barrier against miscarriages of justice, it is indeed necessary to make a breakthrough in this area by ensuring that a previous conviction can be overturned even if evidence emerges that raises a reasonable doubt as to the guilt of the previously convicted person. Moreover, if the imposition of a penalty is subject to a certain standard of proof and it is established, albeit after the final conviction, that the required threshold has not been properly met, the maintenance of the penalty appears to be contrary to an elementary principle of justice.

Finally, it should be noted that in modern “algorithmic societies” (Balkin, 2017, p. 1219), there is a final category of rules that should be revitalised. I refer to the introduction of some criteria for the evaluation of evidence, constructed as negative rules of corroboration, aimed at reducing the risk of judicial overestimation of certain pieces of evidence²²¹.

In my view, such rules could be particularly useful in achieving one objective: to prevent fact-finders from convicting people solely on the basis of elements gathered by artificial intelligence tools, which will inevitably be used more and more in the near future²²². It should be noted that there is a specific reason that justifies this: the adoption of corroboration rules of this kind would in fact constitute an antidote to the risk that, even if the machines act as an aid to the human decision-maker, the latter ends up passively following their indications, since it is difficult to find reasons to distrust them and then to depart from them (Tuzet, 2020, p. 45 ff.). And it is precisely because, faced with this type of evidence, even the rationalising capacity of the instrument of the motivation of judgments is inevitably reduced, that the introduction of targeted corroboration rules could prove valuable. In particular, without prejudice to the prohibition, already enshrined in some European laws²²³, of fully attributing to machines the taking of decisions that may adversely affect the rights of individuals, one could, in particular, codify some well-chosen negative proof rules of evidence, according to which the human agent must always verify that the output of investigative and/or evidentiary tools based on algorithms (think, for example, of the match of a facial recognition tool or blockchain analysis software) is consistent with

to be acquitted in the context of a revision, it is sufficient that the new evidence is capable of raising even a reasonable doubt. See the judgment of the Court of Cassation of 12 May 2004, Contena, p. 679. Similar problems have also arisen in France as a result of a restrictive interpretation of the concept of “doubt” for the purposes of granting an application for revision. On this point, see Verhesschen and Fijnaut (2020, p. 23 f.), who analyse the reform adopted by the French legislature in 2014 to remedy this situation.

²²¹ In this respect, it should be noted that in recent years, a body of thought has developed in both common law and civil law systems to the effect that the time is ripe to establish some criteria for the assessment of evidence that are capable of better defining the limits of judicial discretion, especially in the case of evidence where there is a greater risk of judicial error. See, for example, Damaška (2019, p. 146 ff.); Lupária (2021, p. 119 f.) and Thaman (2016, p. 108).

²²² On the impact that algorithms and AI are already having on criminal justice, see Quattrococo (2020).

²²³ See, for instance, Article 11 of Directive EU 2016/680 (the so-called LED Directive).

other evidence²²⁴. In short, the idea is to create rules that ensure that AI evidence cannot stand alone as “unique and decisive evidence”.

It should be added that the approach thus outlined should be even more rigorous in the event that the individual AI tool presents transparency problems²²⁵ and acts like a black box, i.e. like a system in which inputs and outputs are observable, while the internal functioning remains obscure even to its own programmers (Contissa, Lasagni and Sartor, 2019, p. 620). In such a case, in order to safeguard the pillars of a fair trial, in its articulation of the right to cross-examination and equality of arms, it would indeed be desirable to act directly upstream at the level of admissibility of evidence by providing that «the court should exclude the “black box evidence” from the assessment of the defendant’s guilt» (Quattrocolo, 2020, p. 96).

In conclusion, the introduction of such rules would not mean adopting a reactionary approach aimed at closing the doors to AI tools, given their great potential in the context of criminal justice. Nor would it mean a return to a system of legal proof in which it is always a network of rules on the value of evidence that determine when the standard of proof for conviction is reached or not²²⁶. Moreover, as we have shown, such systems run the risk of being too rigid to function properly and are therefore open to distorting practices. On the contrary, it would only be a question of introducing a limited number of exceptions to the general principle of free evaluation of evidence, which, on closer examination, would be motivated by the desire to safeguard its most profound meaning: that of ensuring that, in the future too, judicial decisions will always be based, at least in part, on rationality and human logic, and not only on that of machines, which is often incomprehensible to us²²⁷.

BIBLIOGRAPHY

Abbagnano (1998). *Dizionario di filosofia*. Utet.

Accatino, D. (2011). Certezas, dudas y propuesta en torno al estándar de la prueba penal. *Revista de Derecho de la Pontificia Universidad Católica de Valparaíso*, 37, p. 483-511.

²²⁴ It is worth noting that some negative rules of this kind have been included in the part of the EU AI law (Regulation (EU) 2024/1689) dealing with remote biometric identification systems. See, for example, Article 5(3), which states: «no decision that produces an adverse legal effect on a person may be taken based solely on the output of the ‘real-time’ remote biometric identification system». See also Article 26 (10) on post-remote biometric identification systems.

²²⁵ The principle of transparency is explicitly recognised in the Ethical Charter for the use of artificial intelligence and other related fields, adopted by the CEPEJ at its XXXI plenary meeting, Strasbourg, 3 December 2018 (CEPEJ, 2018, p. 14).

²²⁶ The prospect of a substantial return to a legal proof system in the near future through artificial intelligence tools has recently been highlighted by Tuzet (2023, p. 244-246).

²²⁷ As Nieva Fenoll (2020, p. 142) recalls «*el Imperio de la Razón fue una conquista social. Merecería la pena no perderlo*».

- Aguilera, E. (2021). Dudando de las dudas sobre la formulación de estándares de prueba jurídicos (razonablemente) objetivos. In C. Vázquez (Ed.), *Ciencia y Justicia. El conocimiento experto en la Suprema Corte de Justicia de la Nación* (pp. 403-452). SCJN.
- Ala, R. (1829). *Pratica criminale* (Vol. II). Giunchi e Comp.
- Ala, R. (1839). *Criminalis juris et praxis instituta usi etiam forensi accommodata* (Vol. I). Mugnoz.
- Alcácer Guirao, R. (2021). Algunas dudas sobre la duda razonable. Prueba de descargo, estándares de prueba e in dubio pro reo. *Revista Electrónica de Ciencia Penal y Criminología*, 23-09, p. 1-47.
- Alessi Palazzolo, G. (1979). *Prova legale e pena. La crisi del sistema tra Evo Medio e moderno*. Jovene.
- Allard, A. (1868). *Histoire de la justice criminelle au seizième siècle*. Libraire-éditeur.
- Allen, R.J. (2013). Los estándares de prueba y los límites del análisis jurídico. In C. Vázquez (Ed.), *Estándares de prueba y prueba científica* (pp. 41-64). Marcial Pons.
- Allen, R.J. (2014). Burdens of proof. *Law, Probability and Risk*, 13(3-4), p. 195-219.
- Allen, R. and Pardo M. (2007). The Problematic Value of Mathematical Models of Evidence. *Journal of Legal Studies*, 36(1), p. 107-140.
- Allen, R. and Pardo M.S (2019). Relative plausibility and its critics. *The International Journal of Evidence & Proof*, 23(1-2), p. 5-59.
- Ambos, K. (2023). “Intime conviction” in Germany. Conceptual foundations, historical development and current meaning. *Quaestio facti*, 4, p. 167-189.
- Anderson, T., Schum D. and Twining, W. (2005). *Analysis of Evidence* (2nd ed.). Cambridge University Press.
- Angles, J. (1587). *Flores theologiarum quaestionum in secundum librum Sententiarum. Pars I*. Stephani Michaelis.
- Antonino da Firenze (1740). *Summa theologica*. Pietro Ballerini.
- Arrivabene, F. (1814). *Della certezza morale nei giudicj penali*. Nicolò Bettoni.
- Barbacovi, F. V. (1820). *Degli argomenti ed indizi nei giudizi criminali*. Società tipografica dei classici italiani.
- Baldo degli Ubaldi (1578). *In Quartum et Quintum Codicis Libr*. Venice.
- Balkin, J.M. (2017). The Three Laws of Robotics in the Age of Big Data. *Ohio State Law Journal*, 78(5), p. 1217-1241.
- Baron (2022). Certainty, in *The Stanford Encyclopaedia of Philosophy*. <https://plato.stanford.edu/entries/certainty>.
- Bartolo da Sassoferrato (1562). *In Ius Universum Civile, Commentaria: Digestum veterum*. Froben.
- Bartolomeo da Saliceto (1578). *In Tertium & Quartum Codicis Libros*. Aquilae.
- Basilico, H. (1691). *Decisiones Criminales*. Thypographia Joannis Philippi Cecchi.
- Beccaria (1880). On Crimes and Punishments. In J.A. Farrer, *Crimes and Punishments. Including a New Translation of Beccaria's “Dei delitti e delle pene”*. Chatto & Windus.
- Benedict XIV (1778). *Sanctissimi domini nostri Benedicti Papae XIV Bullarium tomus tertius: in quo continentur constitutiones, epistolae aliaque edita ab exitu anni MDCCXLVIII usque ad totum pontificatus duodecimum seu ad annum MDCCLII: cum appendice et supplement*. Ioannis Gatti.
- Bentham, J. (1827). *Rationale of Judicial Evidence* (Vol. I). Hunt and Clarke.
- Best, W.M. (1849). *A Treatise on the Principles of Evidence and Practice as to Proof in Court of Common Law*. Sweet.
- Bianchi, M. (1554). Tractatus de indiciis homicidii ex proposito comissi & de alijs indicijs homicidij et furti ad legem finalem ff. de quaestionibus. In *Tractatus Universi Iuris, Duce, et Auspice Gregorius XIII, in unum congesti* (Vol. XI, Part. I, 260). Venice.
- Biel, G. (1574). *Commentarius in Quartum Librum Sententiarum*, Bozolan.
- Bossi, E. (1566). *Tractatus varii*, Jacobi Junctae.
- Blackstone, W. (1770). *Commentaries on the Laws of England* (Vol. IV). Clarendon Press.
- Briganti, T. (1842). *Pratica criminale* (Vol. I). Gabriele Marotta.
- Broun, K. S., Kaye, D. H., Dix, G. E., Imwinkelried, E. J., Mosteller, R. P., Roberts, E. F. and Swift, E. (2014). *McCormick on Evidence* (7th ed.). West Academic Publishing.

- Brugnoli, G. (1846). *Della certezza e prova criminale col confronto di varie legislazioni d'Europa ed in specie d'Italia*. Carlo Vincenzi.
- Caprioli, F. (2009). L'accertamento della responsabilità penale "oltre ogni ragionevole dubbio". *Rivista italiana di diritto e procedura penale*, 52(1), p. 51-92.
- Caprioli, F. (2017). Verità e giustificazione nel processo penale. *Revista Brasileira de Direito Processual Penal*, 3(1), p. 317-341.
- Carlizzi, G. (2018). *Libero convincimento e ragionevole dubbio nel processo penale*. Bonomo editore.
- Carmignani, G. (1852). *Scritti inediti* (Vol. IV). Giuseppe Giusti.
- Caramuel Lobkowitz, I. (1675). *Theologiae moralis fundamentalis* (Lib. IV). Officina Anissoniana.
- Carpzov, B. (1684). *Practica nova imperialis saxonica rerum criminalium*, Balthas. Christophori Wustii.
- Carnelutti, F. (1963). *Verso la riforma del processo penale*. Morano.
- Catalano, E.M. (2016). *Ragionevole dubbio e logica della decisione*. Giuffrè.
- Cherubini, P. (2024). *Guida al ragionamento probatorio con le reti bayesiane*. Giuffrè.
- Claro, G. (1586). *Opera omnia sive practica civilis atque criminalis*. Augustae Taurinorum.
- Clermont, K.M. (2013). *Standard of Decision in Law*. Carolina Academic Press.
- Cohen, L.J. (1977). *The Probable and the Provable*. Clarendon Press.
- Cohen., L.J. (1998). *Introduzione alla filosofia dell'induzione e della probabilità*. Giuffrè.
- Concina, D. (1761). *Della storia del probabilismo e del rigorismo*. Simone Occhi.
- Conti, C. (2020). Il BARD paradigma di metodo: legalizzare il convincimento senza riduzionismi aritmetici. *Diritto penale e processo*, 6, p. 829-842.
- Contigiani, N. (2007). Il processo penale pontificio. Tra ancoraggi inquisitori e spettro riformista (1831-1858). *Rivista di storia del diritto italiano*, 80, p. 189-314.
- Contissa, G., Lasagni, G., Sartor, G. (2019). Quando a decidere in materia penale sono (anche) algoritmi e IA: alla ricerca di un rimedio effettivo. *Diritto di internet*, 4, p. 619-634.
- Cordero, F. (2012). *Procedura penale* (9th ed.). Giuffrè.
- Dahlman C. (2018). Determining the base rate for guilt. *Law, Probability and Risk*, 17(1), p. 15-28.
- Dahlman C. (2020). De-Biasing Legal Fact-Finders With Bayesian Thinking. *Topics in Cognitive Science*, 12(4), p. 1115-1131.
- Dahlman C. (2024). What does it mean to be «plausible»? *Quaestio facti*, 7, p. 91-102.
- Dahlman, C. and Kolfaath, E. (2021). The Problem of the Prior in Criminal Trials. In C. Dahlman, A. Stein and G. Tuzet (Eds.), *Philosophical Foundations of Evidence Law* (pp. 287-300). Oxford University Press.
- Damaška, M. (2019). *Evaluation of Evidence. Pre-Modern and Modern Approaches*. Cambridge University Press.
- Damaška, M. (2022). The future of free proof in criminal cases. In P. Bárd (Ed.). *Liber Amicorum Karoly Bard* (pp. 197-204, Vol. II). L'Harmattan.
- Daniele, M. (2009). *Regole di esclusione e regole di valutazione della prova*. Giappichelli.
- Deciani, T. (1613). *Tractatus criminalis* (Vol. I). Ioannis Bringeri.
- De Coras (1668). *In aliquot titulos & capita legum Secundi ac tertii tomi Pandectarum*. Antonium Vicentium.
- De Damhouder, J. (1556). *Praxis rerum criminalium*. Antwerp.
- Dei Vecchi, D. (2022). Prueba sin convicción en su justa medida. *Doxa*, 45, p. 337-373.
- Dalia, G. (2018). *Convincimento giudiziale e ragionevole dubbio*. Wolters Kluwer-Cedam.
- Della Torre, J. (2014). Il lungo cammino della giurisprudenza italiana sull'«oltre ogni ragionevole dubbio». *Diritto penale contemporaneo*. <https://archiviopdc.dirittopenaleuomo.org/d/3161-il-lungo-cammino-della-giurisprudenza-italiana-sull-oltre-ogni-ragionevole-dubbio>
- Della Torre, J. (2015). Standard di prova e condanna penale: una ricostruzione metateorica e metagiurisprudenziale. *Materiali per una storia della cultura giuridica*, 45(2), p. 369-492.
- Della Torre, J. (2019). *La giustizia penale negoziata in Europa. Miti, realtà e prospettive*. Wolter-Kluwer.
- Della Torre, J. (2023a). Gli standard probatori per la condanna penale tra storia e attualità. *La legislazione penale*. <https://www.la legislazione penale.eu/wp-content/uploads/2023/11/Della-Torre-LP-1.pdf>

- Della Torre, J. (2023b). Standard probatori, efficienza e giustizia penale. *Politica del diritto*, 54(3), pp. 327-364.
- Della Torre, J. (2024). La ragionevole previsione di condanna alla prova della realtà giudiziaria. *La legislazione penale*. https://www.lalegislazionepenale.eu/wp-content/uploads/2024/07/Della-Torre_LP_.pdf.
- Delmas-Marty, M. (1996). La preuve pénale. *Droits*, 23, p. 53-65.
- De Lugo, J. (1696). *Disputationes Scholasticae et morales*. Petri Borde, Joan & Petri Arnaud.
- De Rosa, C.A. (1680). *Decretorum M.C. Praxis Criminalis*. Hyacinthum Passarum.
- Descartes, R. (1667). *Les Principes de la Philosophie*. Henry Le Gras.
- Dezza, E. (2013). *Lezioni di storia del processo penale*. Pavia University Press.
- Dripps, D.A. (1987). The Constitutional Status of the Reasonable Doubt Rule. *California Law Review*, 75, p. 1665-1718.
- Duce, M. (2024). Plea bargaining and the risk of wrongful convictions: a comparative overview. In M. Langer et al. (eds.), *Research Handbook on Plea Bargaining and Criminal Justice* (pp. 278-297). Elgar.
- Ede, R. and Ford, C. (2004). *The Law Society, Study of The Laws of Evidence in Criminal Proceedings throughout the European Union, Summary Report*. <https://www.scribd.com/document/144338382/The-Law-Society-Study-of-the-laws-of-evidence-in-criminal-proceedings-throughout-the-European-Union-Summary-ReportOctober-2004>
- Eggleston, R. (2004). *Prova, conclusione probatoria e probabilità* (2nd ed.). Giuffrè.
- Esmein, A. (1882). *Histoire de la procédure criminelle en France et spécialement de la procédure inquisitoire, depuis le XVIIe siècle jusqu'à nos jours*. Libraires - éditeurs.
- Esnard, C., Grihom, M. J. and Leturmy, L. (2015). *L'intime conviction : incidences sur le jugement des jurés et magistrates. Rapport Final. 8 Juillet 2015*. <http://www.gip-recherche-justice.fr/wp-content/uploads/2016/01/Rapport-Final-Intime-Conviction.pdf>.
- European Commission for the Efficiency of Justice [CEPEJ] (2018). *European ethical Charter on the use of Artificial Intelligence in judicial systems and their environment*. <https://rm.coe.int/ethical-charter-en-for-publication-4-december-2018/16808f699c>
- European Court of Human Rights [ECtHR] (1950). *European Convention on Human Rights*. https://www.echr.coe.int/documents/d/echr/convention_ENG
- Farinacci, P. (1607). *Consilia sive responsa atque Decisiones causarum Criminalium*. Horatij Cardon.
- Farinacci, P. (1616). *Praxis et theoricæ criminalis*, Horatij Cardon.
- Ferrer Beltrán, J. (2004). *Prova e verità nel diritto*. Il Mulino.
- Ferrer Beltrán, J. (2007). *La valutazione razionale della prova*. Giuffrè.
- Ferrer Beltrán, J. (2021a). Proven Facts, Beliefs, and Reasoned Verdicts. In C. Dahlman, A. Stein and G. Tuzet (eds.), *Philosophical Foundations of Evidence Law* (pp. 40-52), Oxford University Press.
- Ferrer Beltrán, J. (2021b). *Prueba sin convicción. Estándares de prueba y debido proceso*. Marcial Pons.
- Ferrer Beltrán, J. and Tuzet, G. (2018). Sulla necessità degli standard di prova per la giustificazione delle decisioni giudiziali. *Diritto & Questioni pubbliche*, 2, p. 455-472.
- Ferrua, P. (2020). Onere della prova, regola di giudizio e standard probatorio. *Cassazione penale*, 60, p. 2639-2649.
- Feuerbach, P.J.A. (1813). *Betrachtungen über das Geschwornen-Gericht*. Philipp Krüll.
- Filangieri, G. (1806). *La scienza della legislazione* (Vol. I). Santini.
- Finklestein M.O. and Fairley W.B. (1970). A Bayesian approach to Identification Evidence. *Harvard Law Review*, 83(3), 489-517.
- Fiorelli, P. (2023). *La tortura giudiziaria nel diritto comune*. Giuffrè.
- Fortunato, S.J. (1996). Instructing on Reasonable Doubt after Victor v. Nebraska: A Trial Judge's Certain Thoughts on Certainty. *Villanova Law Review*, 41, p. 365-431.
- Franklin, J. (2001). *The Science of Conjecture. Evidence and Probability before Pascal*. John Hopkins University.
- Friedman, R.D. (1997). Answering the Bayesioskeptical Challenge. *The International Journal of Evidence and Proof*, 1, p. 276-291.
- Friedman, R.D. (2000). A Presumption of Innocence, Not of Even Odds. *Stanford Law Review*, 52(4), p. 873-587.

- Gail, A. (1613). *Observationum practicarum Imperialis Camerae*. Petrum de Farris.
- Garbolino (2014). *Probabilità e logica della prova*. Giuffrè.
- Garlati, L. (2004). Il diabolico intreccio. Reo convinto e indizi indubitati nel commento di Bartolomeo da Saliceto (C. 4.19.25): alle radici di un problema. In O. Condorelli (ed.), *Panta rei. Studi dedicati a Manlio Bellomo* (pp. 387-419, Vol. II). Il Cigno Edizioni.
- Gerson, J. (1706). *Joannis Gersoni Opera Omnia* (t. I). Sumptibus Societatis.
- Giuliani, G. (1840). *Istituzioni di diritto criminale col commento della legislazione gregoriana*. Viarchi.
- Glaser (1883). *Beiträge zur Lehre vom Bewies im Strafprozess*. Duncker und Humblot.
- González Lagier, L. (2020). ¿Es posible formular un estándar de prueba preciso y objetivo? Algunas dudas desde un enfoque argumentativo de la prueba. *Revista Telemática de Filosofía del Derecho*, 23, p. 79-97.
- Greenleaf, S. (1842). *A Treatise on the Law of Evidence* (Vol. I). Charles C. Little & James Brown.
- Guérin, M.C. (2015). Intime conviction et doute. Au-delà d'une antinomie apparente...Ce que révèle leur étude comparée. *Annales Médico-psychologiques*, 173, p. 610-614.
- Haack, S. (2014). *Evidence Matters, Science, Proof, and Truth in the Law*. Cambridge University Press.
- Hale, M. (1778). *History of the Pleas of the Crown* (Vol. I). George Wilson.
- Hamer, D. (2007). The Presumption of Innocence and Reverse Burdens: A Balancing Act. *Cambridge Law Journal*, 66(1), p. 142-171.
- Hélie, F. (1867). *Traité de l'instruction criminelle* (Vol. VIII). Henri Plon.
- Hélie, F. (1870). *Des délits et des peines par Beccaria. Deuxième édition. Avec une introduction et un commentaire Par M. Fausin Hélie*. Gaillaumin.
- Ho, H.L. (2008). *A Philosophy of Evidence Law. Justice in the Search for Truth*. Oxford University Press.
- Human Rights Committee. United Nations (2007). *General comment no. 32. Article 14: right to equality before courts and tribunals and to a fair trial*. Doc. CCPR/C/GC/32 of 23 August 2007. https://tbinternet.ohchr.org/_layouts/15/treatybodyexternal/Download.aspx?symbolno=CCPR%2FC%2FGC%2F32&Lang=en
- Iacoviello, F.M. (2006). Lo "standard probatorio dell'"al di là di ogni ragionevole dubbio" e il suo controllo in cassazione. *Cassazione penale*, p. 3869-3884.
- Iacoviello, F.M. (2023). *La Cassazione penale. Fatto, diritto e motivazione* (2nd ed.). Giuffrè.
- Incampo, A. and Scafati, A. (eds.) (2017). *Giudizio penale e ragionevole dubbio*. Cacucci.
- International Criminal Court [ICC] (2021). *Rome Statute of the International Criminal Court*. <https://www.icc-cpi.int/sites/default/files/2024-05/Rome-Statute-eng.pdf>
- Jellema, H. (2023). *(Im)probable stories: combining Bayesian and explanation-based accounts of rational criminal proof*. University of Groningen. <https://doi.org/10.33612/diss.563508244>.
- Judicial College (2023). *The Crown Court Compendium. Part I: Jury and Trial Management and Summing Up*. <https://www.judiciary.uk/wp-content/uploads/2024/09/Crown-Court-Compendium-Part-I-July-2024.pdf>
- Kadane J. and Schum D.A. (1996), *A Probabilistic Analysis of the Sacco and Vanzetti Evidence*. Wiley.
- Kaye, D.H. (1999). Clarifying the burden of persuasion: what Bayesian decision rules do and do not do. *The International Journal of Evidence & Proof*, 3(1), p. 1-28.
- Keane, A and McKeown, P. (2019). Time to Abandon "Beyond Reasonable Doubt" and "Sure": The Case for a New Direction on the Criminal Standard and how it should be used. *Criminal Law Review* 2019, 50(6), p. 505-527.
- Langbein, J. (1974). *Prosecuting Crime in the Renaissance. England, Germany, France*. Harvard University Press.
- Langbein, J. (2003). *The Origins of Adversary Criminal Trial*. Oxford University Press.
- Laudan, L. (2005). Por qué un estándar de prueba subjetivo y ambiguo no es un estándar. *Doxa*, 28, p. 95-113.
- Laudan, L. (2006). *Truth, Error, and Criminal Law*. Cambridge University Press.
- Le Brun de La Rochette, C. (1611). *Le Procès criminel divisé en deux livres* (Vol. II). Pierre Calles.
- Leclerc, H. (1995). L'intime conviction du juge: norme démocratique de la preuve. *Le for intérieur*, p. 206-213.

- Lempert, R. (2003). La nuova dottrina delle prove: un'analisi della formazione della prova concludente. In P. Tillers and E.D. Green (eds.), *L'inferenza probabilistica nel diritto delle prove. Usi e limiti del bayesianesimo* (pp. 91-155). Giuffrè.
- Levy, J. P. (1939). *La Hiérarchie des preuves dans le droit savant du Moyen-Age, depuis la Renaissance du droit romain jusqu'à la fin du XIVe siècle*. Librairie du Recueil Sirey.
- Lillquist, E. (2005). Absolute Certainty and the Death Penalty. *American Criminal Law Review*, 42(1), p. 45-91.
- Lucchini, L. (1895). *Elementi di procedura penale*. Barbera.
- Lupária, L. (2021). Notazioni controintuitive su intelligenza artificiale e libero convincimento. In *Intelligenza artificiale, giurisdizione penale ed etica del giudizio* (pp. 113-122). Giuffrè.
- Makor, A.R. (2024). Risks of Incorrent Use of Probabilities in Court and What to Do about Them. In A. Placani and S. Broadhead (eds.), *Risk and Responsibility in Context* (pp. 99-108). Routledge.
- Mackor, A.R., Jellema, H. and van Koppen, P.L. (2021). Explanation-Based Approaches to Reasoning about Evidence and Proof in Criminal Trials. In B. Brozek, J. Hage and N. Vincent (eds.), *Law and Mind: A Survey of Law and the Cognitive Sciences* (pp. 431-470). Cambridge University Press.
- Martinon, J. (1663). *Disputationes theologicae* (Vol. II). Sebastianum Cramoisy.
- Mascardi, G. (1585). *De probationibus* (Vol. I, Sif). Feyerabendij.
- Mavidal, M.J. and Laurent, M.E. (Eds.) (1885), *Archives Parlementaires de 1787 a 1860, Recueil Complet des Débats Législatifs et Politiques des Chambres Françaises, Première Série* (Vol. XXII). Paul Dupont.
- Mazza, O. (2012). Il ragionevole dubbio nella toeria della decisione. *Criminalia*, p. 357-371.
- Mazza, O. (2025). Distopia del processo artificiale. *Archivio penale web*, 1. <https://archiviopenale.it/File/DownloadArticolo?codice=36e0a0e4-7576-4379-a107-a31292613919&idarticolo=46845>
- McKeown, P. (2022). Dreaded questions, doubtful answers & the trouble with 'sure'. *NewLawJournal*, 7989. <https://www.newlawjournal.co.uk/content/dreaded-questions-doubtful-answers-the-trouble-with-sure>
- Medina, F. B. (1580). *Expositio in prima secundae Angelici Doctoris d. Thomae Aquinatis*. Petrum Dehuchinum.
- Mittermaier, C. J. A. (1834). *Die Lehre vom Beweise im deutschen Strafprozesse nach Fortbildung durch Gerichtsgebrauch und deutsche Gesetzbücher in Vergleichung mit den Ansichten des englischen und französischen Strafverfahrens*. Johann Wilhelm Heyer.
- Mittermaier, C. J. A. (1850). *Trattato della pruova in materia penale*. Giuseppe Carluccio.
- Mittermaier, C. J. A. (1851). *Il processo orale accusatorio e per giurati secondo le varie legislazioni*. Stefano Calderini e comp. – Nicola Zanichelli e comp.
- Mittermaier, C. J. A. (1858). *Teoria della prova nel processo penale*. Francesco Sanvito.
- Morano, A. A. (1975). A re-examination of the Development of the Reasonable Doubt Rule. *Boston University Law Review*, 55(3), p. 507-529.
- Nance, D. A. (2016). *The Burden of Proof. Discriminatory Power, Weight of Evidence, and Tenacity of Belief*. Cambridge University Press.
- Neri, R. (2024). Judging Beyond Any Reasonable Doubt: A Logic and Epistemological Rule. *Quaestio Facti*, 7, p. 43-68.
- Nieva Fenoll, J. (2010). *La valoración de la prueba*. Marcial Pons.
- Nieva Fenoll, F. (2020). Carga de la prueba y estándares de prueba: dos reminiscencias del pasado. *Estudios de Derecho*, p. 119-148. <https://doi.org/10.17533/udea.esde.v77n170a05>
- Nobili, M. (1974). *Il principio deli libero convincimento del giudice*. Giuffrè.
- Nobili, M. (2004). Esiti, errori, arbitrii dietro un'illustre formula: gli ultimi trent'anni. In *Il libero convincimento del giudice penale. Vecchie e nuove esperienze* (33-59). Giuffrè.
- Padoa-Schioppa, A. (1999). Sur la conscience du juge dans le ius commune. In J.-M. Carbasse and L. Depambour-Tarride (Eds.), *La Conscience du juge dans la tradition juridique européenne* (pp. 95-129). Press Universitaires De France.
- Pagano, F. M. (1824). *Principj del codice penale e logica de' probabili*. Luca Marotta.
- Pardo, M. S. (2023). What Makes Evidence Sufficient?. *Arizona Law Review*, 65, p. 431-478.

- Patuzzi, J. V. (1790). *Ethica christiana sive Theologia moralis* (Vol. I). Remondini.
- Penim Pinheiro, J. (2021). Princípio in dubio pro reo - considerações gerais. *Julgar Online*. <https://julgar.pt/principio-in-dubio-pro-reo-consideracoes-gerais/>.
- Picinali, F. (2012). Structuring Inferential Reasoning in Criminal Fact Finding: An Analogical Theory. *Law, Probability and Risk*, 11(2-3), p. 197-223.
- Picinali, F. (2015). The Threshold Lies in the Method: Instructing jurors about reasoning beyond reasonable doubt. *The International Journal of Evidence & Proof*, 19(3), p. 139-153.
- Picinali, F. (2022). *Justice In-Between. A Study of Intermediate Criminal Verdicts*. Oxford University Press.
- Pierro, G. (2011). *Accertamento del fatto e colpevolezza al di oltre di ogni ragionevole dubbio*. Aracne.
- Poli, R. (2023). *Prova e convincimento giudiziale del fatto*. Giappichelli.
- Poullain-Duparc, A.M. (1771). *Principe du droit français, suivant les maximes de Bretagne* (Vol. XI). François Avatar.
- Pradel J. (1985). Le déroulement du procès pénal français (aperçus comparatifs avec le droit canadien). *Revue générale de droit*, 16(3), p. 575-590.
- Pradel J. (2016). *Droit pénal comparé* (4th ed.). Dalloz.
- Prakken, H., Bex, F., Mackor, A. R. (2020). Models of Rational Proof in Criminal Law. *Topics in Cognitive Science*, 12(4), p. 1053-1067.
- Pufendorf, S. (1759). *De Jure Naturæ et Gentium*. Officina Knochiana.
- Quattrococo, S. (2020). *Artificial Intelligence, Computational Modelling and Criminal Proceedings. A Framework for a European Legal Discussion*. Springer.
- Raynaldi, G. D. (1735). *Obsevationum criminalium, civilium, & Mixtarum, Liber II, Observ.* Typographia Balleoniana.
- Richard, J. (2017). *L'intime conviction du juge en matière criminelle*. [Doctoral Thesis. Université Montpellier] https://theses.hal.science/tel-01610334v1/file/65991_RICHARD_2017_archivage.pdf.
- Roberts, P. (2011). Does Article 6 of the European Convention on Human Rights Require Reasoned Verdicts in Criminal Trials?. *Human Rights Law Review*, 11(2), p. 213-235.
- Roberts, P. (2014). Loss of Innocence in Common Law Presumptions. *Criminal Law and Philosophy*, 8, p. 317-336.
- Roberts, P. and Zuckerman, A. (2022). *Criminal Evidence* (3th ed.). Oxford University Press.
- Romagnosi, G.D. (1874). *Cenni sui sistemi della convinzione intima e delle prove legali, e sul calcolo degli indizi*. In *Opere di G.D. Romagnosi* (Vol. VII, pp. 665-669). Giuseppe Pedone – Giuseppe Marghieri.
- Rosoni (1995). *Quae singula non prosunt collecta iuvant. La teoria della prova indiziaria nell'età medievale e moderna*. Giuffrè.
- Rovito, S. (1634). *Decisiones Supremorum Tribunalium Regni Neapolitani*. Iacobum Gassarum.
- Saltzburg, S. A. and Capra, D.J. (2018). *American Criminal Procedure. Cases and Commentary* (11th ed.). West Academic Publishing.
- Sanchez, T. (1622). *Operis moralis in praecepta decalogi* (Vol. II). Martinum Nutium.
- Saraceno P. (1940). *La decisione sul fatto incerto nel processo penale*. Cedam.
- Schum D. A. (1994). *The Evidential Foundations of Probabilistic Reasoning*. Northwestern University Press.
- Shapiro, B. J. (1991). *Beyond Reasonable Doubt and Probable Cause*. University of California Press.
- Shapiro, B. J. (2009). Changing Language, Unchanging Standard: From “Satisfied Conscience” to “Moral Certainty” and “Beyond Reasonable Doubt”. *Cardozo Journal of International and Comparative Law*, 17(2), p. 261-280.
- Shapiro, B. J. (2012). Beyond Reasonable Doubt: The Evolution of a Concept. In Y. Batsaki et al. (Eds.), *Fictions of Knowledge* (pp. 19-39). Palgrave Macmillan.
- Shapiro, B. J. (2014). ‘Beyond Reasonable Doubt’: The Neglected Eighteenth-Century Context. *Law and Humanities*, 8(1), p. 19-52.
- Shapiro, B. J. (2022). The Problem of Certainty and the Changing Status of Probable Proofs. *Philosophical Inquiries*, 10(2), p. 117-137.

- Shapiro, J.A. and Muth K.T. (2021). Al di là di ogni ragionevole dubbio: le giurie non lo capiscono. *Criminalia*, p. 53-69.
- Sheppard, S. (2003). The Metamorphoses of Reasonable Doubt: How Changes in the Burden of Proof Have Weakened the Presumption of Innocence. *Notre Dame Law Review*, 78(4), p. 1165-1249.
- Shüssler, R. (2009). Jean Gerson, Moral Certainty and the Renaissance of Ancient Scepticism. *Renaissance Studies*, 23, p. 445-462.
- Shüssler, R. (2019). *The Debate on Probable Opinions in the Scholastic Tradition*. Brill.
- Sinistrari (1754). *Formularium criminale. Pars II*. Caroli Giannini.
- Smith (1748). *The Britons and Saxons Not Converted to Popery*. James Bettenham.
- Smith, O.K.H. (2022). The beyond a reasonable doubt standard of proof: Juror understanding and reform. *The International Journal of Evidence & Proof*, 26(4), p. 291-308.
- Solan, L.M. (1999). Refocusing the Burden of Proof in Criminal Cases: Some Doubt about Reasonable Doubt. *Texas Law Review*, 78, p. 105-147.
- Starkie, T. (1826). *A Practical Treatise of the Law of Evidence, and Digest of Proofs, in Civil and Criminal Proceedings* (Vol. I). West and Lilly.
- Stella, F. (2003). *Giustizia e modernità. La protezione dell'innocente e la tutela delle vittime* (3th ed.). Giuffrè.
- Stein, A. (2005). *Foundations of Evidence Law*. Oxford University Press.
- Suarez, F. (1741). *De Divina Gratia* (Vol. VIII). Balleoniana.
- Summers, S.J. (2023). The Epistemic Ambitions of the Criminal Trial: Truth, Proof, and Rights. *Quaestio facti*, 4, p. 249-272.
- Stephen (1883). *A History of the Criminal Law of England* (Vol. I). MacMillan.
- Taroni, F., Aitken, C., Garbolino, P., Biedermann, A. (2006). *Bayesian Networks and Probabilistic Inference in Forensic Science* (2nd ed.). Wiley.
- Taroni, F., Biedermann, A. and Aitken, C. (2023). Statistical Interpretation of Evidence: Bayesian Analysis. In M.M. Houck (Ed.), *Encyclopedia of Forensic Sciences* (3rd ed., pp. 656-663). Elsevier.
- Taruffo, M. (1992). *La prova dei fatti giuridici*. Giuffrè.
- Taruffo, M. (2003). Rethinking the Standards of Proof. *The American Journal of Comparative Law*, 51(3), p. 659-677.
- Taruffo, M. (2005). Tres observaciones sobre “Por qué un estándar de prueba subjetivo y ambiguo no es un estándar”, de Larry Laudan. *Doxa*, 28, p. 115-126.
- Taruffo, M. (2009a). Fatto, prova e verità (alla luce del principio dell'oltre ogni ragionevole dubbio). *Criminalia*, p. 306-319.
- Taruffo, M. (2009b). *La semplice verità. Il giudice e la costruzione dei fatti*. Laterza.
- Thaman, S.C. (2016). Ensuring the factual reliability of criminal convictions: reasoned judgments or a return to formal rules of evidence?. In J.E. Ross and S.C. Thaman (Ed.) *Comparative Criminal Procedure* (p. 75-114). Elgar.
- Thomas Aquinas (1265-1274). *Summa Theologiae*.
- Tillers, P. (2011). Trial by mathematics – reconsidered. *Law, Probability and Risk*, 10, p. 167-173.
- Tillers, P. and Gottfried, J. (2006). Case Comment – *United States v. Copeland*, 369 F. Supp. 2d 275 (E.D.N.Y. 2005). *Law, Probability and Risk*, 5, p. 135-157.
- Tribe, L.H. (1971). Trial by Mathematics: Precision and Ritual in the Legal Process. *Harvard Law Review*, 84(6), p. 1329-1392.
- Triggiani, N. (2017). Lo standard dimostrativo nella decisione di colpevolezza. In A. Incampo and A. Scalfati (Eds.), *Giudizio penale e ragionevole dubbio* (pp. 325-338). Cacucci.
- Tuzet, G. (2020). L'algoritmo come pastore del giudice? Diritto, tecnologia, prova scientifica. *Media Laws*, p. 45-55.
- Tuzet, G. (2021). Assessment and Standards of Proof: a Messy Issue. *Quaestio facti*, (2), p. 87-113.
- Tuzet, G. (2023). *La prova ragionata*. Giuffrè.
- Tuzet, G. (2024). Certainty Beyond a Reasonable Doubt. *Contemporary Pragmatism*, 20(4), p. 398-423.
- Twining, W. (1991). The New Evidence Scholarship. *Cardozo Law Review*, 13, p. 295-302.

- Ubertis, G. (2013). Razionalismo processuale e verità. *Il giusto processo civile*, p. 325-333.
- Ubertis, G. (2021). *Profili di epistemologia giudiziaria* (2nd Ed.). Giuffrè.
- Vassalli, G. (1968). Il diritto alla prova nel processo penale. *Rivista italiana di diritto e procedura penale*, p. 3-59.
- Verhesschen, K. and Fijnaut, C. (2020). Correcting Wrongful Convictions in France. *Erasmus Law Review*, 4, p. 22-32.
- Verricelli, A. M. (1653). *Quaestiones morales, & legales*, Franciscum Baba.
- Voltaire (1818) *Ouvres complètes. Politique et Législation* (Vol. I, T. XXI). Plancher.
- Waldron, J. (1994). Vagueness in Law and Language: Some Philosophical Issues. *California Law Review*, 82(3), p. 509-540.
- Welch R.E. (2013). Give Me That Old Time Religion: The Persistence of the Webster Reasonable Doubt Instruction and the Need to Abandon It. *New England Law Review*, 48, p. 31-54.
- Wharton, F. (1880). *A Treatise on the Law of Evidence in Criminal Issues* (8th ed.). Kay and Brother.
- Wigmore, J. (1940). *A Treatise on the System of Evidence in Trials at Common Law* (Vol. 9). Little, Brown & Company.
- Whitman, J.Q. (2008). *The Origins of Reasonable Doubt. Theological Roots of the Criminal Trial*. Yale University Press.
- Zander, M. (2000). The criminal standard of proof: how sure is sure?. *New Law Journal*, 1517-1519.
- Zander, M. (2020). The criminal standard of proof: how sure is sure? Pt 2. *New Law Journal*, 7888. <https://www.newlawjournal.co.uk/content/the-criminal-standard-of-proof-how-sure-is-sure-pt-2>

Legislation and case law

- Ajdarić v. Croatia*, n° 20883/09, European Court of Human Rights, 13 December 2011
- Barberà, Messegué and Jabardo v. Spain*, n° 10590/83, European Court of Human Rights, 7 December 1988
- Boaden, R v* (2019) EWCA Crim 2284
- Bogdanovic, R, v* (2020) EWCA Crim 1229
- Bundesgerichtshof (Germany), 1 StR 360/16, January 12, 2017
- Bundesgerichtshof (Germany), 4 StR 554/16, May 11, 2017
- Bundesgerichtshof (Germany), 1 StR 535/16, July 12, 2017
- Bundesgerichtshof (Germany), 1 StR 261/17, November 21, 2017
- Bundesgerichtshof (Germany), 2 StR 399/21, February 16, 2022
- Cage v. Louisiana*, 498 U.S. 39 (1990)
- Coffin v. United States*, 156 U.S. 453 (1895)
- Commonwealth vs. John W. Webster* (1850) 5 Cush. 295, 59 (Mass. 1850)
- Court of Cassation Sentence (Italy), n. 1368, October 6, 1883. *La Corte Suprema di Roma*, 1883
- Court of Cassation Sentence (Italy), Pressi, February 24, 1981. *Rivista penale*, 1981
- Court of Cassation Sentence (Italy), Mattia, October 22, 1984. *CED Cassazione*, n. 167449
- Court of Cassation Sentence (Italy), Joint Chambers, Franzese, July 10, 2002. *Rivista italiana di diritto e procedura penale*, 2002
- Court of Cassation Sentence (Italy), Contena, May 12, 2004. *Rivista italiana di diritto e procedura penale*, 2004
- Court of Cassation Sentence (Italy), n. 41110, October 24, 2011. *CED Cassazione*, n. 251507
- Court of Cassation Sentence (Italy), n. 42056, May 26, 2016. *DeJure*.
- Court of Cassation Sentence (Italy), n. 48541, June 19, 2018. *CED Cassazione*, n. 274358-01
- Court of Cassation Sentence (Italy), n. 39777, October 2, 2023. *DeJure*
- Court of Milan Sentence (Italy), of June 18, 2015. *Diritto penale contemporaneo*. <https://archiviodpc.dirittopenaleuomo.org/dl/4213-il-teorema-di-bayes-fa-capolino-al-tribunale-di-milano>
- Desir* [2022] EWCA Crim 1071, 198

- Directive (EU) 2016/343 of the European Parliament and of the Council, of 9 March 2016, on the strengthening of certain aspects of the presumption of innocence and of the right to be present at the trial in criminal proceedings (OJ L, num. 65, 11-3-2016, p. 1–11). <https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX%3A32016L0343>
- Directive (EU) 2016/680 of the European Parliament and of the Council, of 27 April 2016, on the protection of natural persons with regard to the processing of personal data by competent authorities for the purposes of the prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties, and on the free movement of such data, and repealing Council Framework Decision 2008/977/JHA (OJ L, num. 119, 4-5-2016, p. 89–131). <https://eur-lex.europa.eu/eli/dir/2016/680/oj/eng>
- Falk v. Netherlands, n° 66273/01, European Court of Human Rights, 19 October 2004
- Gäfgen v Germany, n° 22978/05, European Court of Human Rights, 3 June 2010
- Hepworth and Fearnley (1955) 2 QB 600, 603
- Holland v. United States, 348 U.S. 140 (1954)
- In re Winship, 397 U.S. 358 (1970)
- Jackson v. Virginia, 443 U.S. 307, 309, 323-24 (1979)
- Kerimoğlu v. Turkey, n° 29600/10, European Court of Human Rights, 6 December 2022
- Kritz [1950] K.B. 82, 88-90
- Lifchus (1997) 3 SCR 320
- Majid [2009] EWCA Crim 2563
- Manuela et al. v. El Salvador, Inter-American Court of Human Rights, November 2, 2021
- Miah (2018) EWCA Crim 563
- Miles v. United States, 103 U.S. 312 (1880)
- Miller v. Minister of Pensions (1947) 2 All ER 372
- Mohammad & Another (2022) EWCA Crim 380
- Murphy v. Holland, (1985) 776 F.2d470, (4th Cir. 1985)
- People v. Downs (2015) IL 117934, 32. (Il. 2015)
- Police and Criminal Evidence Act 1984, c. 60 <https://www.legislation.gov.uk/ukpga/1984/60/contents>
- R v JL (2017) EWCA Crim 621
- R v Yap Chuan Ching (1976) 63 Cr App Rep 7, 11
- Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act) (Text with EEA relevance) (OJ L, num. 2024/1689, 12-7-2024). <https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng>
- Ramos v. Louisiana, 590 U.S. (2020)
- Roura v. The State, 214 Ga. App. 43 (Ga. 1994)
- Salabiaku v. France, n° 10519/83, European Court of Human Rights, 7 October 1988
- Sullivan v. Louisiana, 508 U.S. 275 (1993)
- Summers (1952) 36 Cr. App. R. 14, 15
- Supremo Tribunal de Justiça Sentence (Portugal), SJ200801100041985, of January 10, 2008. <https://www.dgsi.pt/jstj.nsf/954f0ce6ad9dd8b980256b5f003fa814/7bfd4fdcd30787d802573d2003c2015?OpenDocument>
- State v. Mitchell, 577 N.W.2d 481, 485 (Minn. 1998)
- Taxquet v. Belgium, n° 926/05, European Court of Human Rights, 16 November 2010
- Tribunal Constitucional Sentence (Spain), 129/1998 (First Chamber), of 16 June 1998 (appeal 1331/1995) - ES:TC:1998:129
- Tribunal Supremo Sentence (Spain), 2254/2023 (Criminal Court), 21 September 2023
- Victor v. Nebraska, 511 U.S. 14 (1994).
- Wanhalla (2006) NZCA 229.
- Woolmington v D.P.P. (1935) AC 462.

CIENCIA
PARA EL PROCESO

FALSAS CREENCIAS SOBRE LA MEMORIA DE LOS TESTIGOS EN PROFESIONALES Y ESTUDIANTES DE CIENCIAS SOCIALES Y JURÍDICAS

Antonio L. Manzanero

Universidad Complutense de Madrid

Álvaro Alanzavez

Policía Nacional

Jordi Ferrer Beltrán

Universidad de Girona

Carmen Vázquez

Universidad de Girona

RESUMEN: Se evaluó a policías con y sin experiencia, jueces en formación en la Escuela Judicial y estudiantes del Grado de Psicología, Derecho y Criminología sobre algunos de los principales tópicos acerca de la psicología del testimonio. Los resultados mostraron un porcentaje medio de aciertos globales cercano al azar, con escasas diferencias entre los distintos colectivos evaluados. Los resultados muestran que son frecuentes las falsas creencias sobre el funcionamiento de la memoria de los testigos, se desconocen los factores que influyen en la exactitud de las declaraciones e identificaciones y la capacidad para testificar de testigos infantiles y de avanzada edad. Se concluye que muchas de las creencias sobre la memoria de los testigos son falsas. Una formación mixta en psicología, derecho y ciencias policiales podría favorecer un mejor conocimiento sobre el funcionamiento de la memoria de los testigos que permita minimizar los errores en la valoración de los testimonios.

PALABRAS CLAVE: valoración de la prueba testifical, memoria de testigos, psicología del testimonio, creencias.

FALSE BELIEFS ABOUT WITNESS MEMORY IN PROFESSIONALS AND STUDENTS OF SOCIAL AND LEGAL SCIENCES

ABSTRACT: Police officers with and without experience, trainee judges at the Judicial School, and undergraduate students in Psychology, Law, and Criminology were assessed on some of the main topics regarding the psychology of testimony. The results showed an average percentage of correct answers close to chance, with few differences between the various evaluated groups. The findings indicate that false beliefs about the functioning of witness memory are common, factors influencing the accuracy of statements and identifications are unknown, and there is limited understanding of the testimony capacity of child and elderly witnesses. It is concluded that many beliefs about witness memory are false. A mixed training in psychology, law, and police sciences could promote a better understanding of witness memory functioning, allowing for the minimization of errors in testimony assessment.

KEYWORDS: testimonial evidence assessment, witness memory, psychology of testimony, beliefs.

SUMARIO: 1. INTRODUCCIÓN.— 2. MÉTODO: 2.1. Participantes. 2.2. Procedimiento.— 3. RESULTADOS: 3.1. Aciertos generales en el cuestionario. 3.2. Policías vs. estudiantes de grado vs. jueces en formación. 3.3. Jueces en formación. 3.4. Policías en función de la experiencia. 3.5. Estudiantes en función de la especialidad.— 4. DISCUSIÓN.— 5. CONCLUSIONES.— BIBLIOGRAFÍA.

1. INTRODUCCIÓN

La prueba testifical es probablemente una de las más importantes, ya que aporta el significado de las demás pruebas, fundamentalmente de la prueba material. Un rastro biológico recogido de la escena del delito, aunque los análisis de ADN y las bases de datos permitan ponerle nombre y apellidos, por sí mismo no dice nada acerca del papel que ha jugado esa persona (víctima, testigo o autor) y ni tan siquiera si estuvo en el lugar donde se cometió el delito antes, durante o después de los hechos. Por otro lado, el derecho a ser escuchado forma parte tanto de los derechos de las víctimas como de los acusados¹.

No obstante, se trata de un tipo de prueba con bajo valor probatorio y que requiere procedimientos específicos (De Paula Ramos, 2019; Nieva Fenoll, 2010; Ramírez Ortiz, 2020; Vázquez y Fernández López, 2022), ya que un testimonio es un recuerdo subjetivo y no siempre exacto de los hechos². Su valoración racional, epistémicamente fundada, sería la principal estrategia de razonamiento, lo que implica aplicar la

¹ Estatuto de la víctima del delito (Ley 4/2015 de 27 de abril). Estatuto de la persona encausada (Anteproyecto de Ley de Enjuiciamiento Criminal de 2020).

² Téngase en cuenta que la memoria es un proceso constructivo, dinámico y en continuo cambio, que se ve afectada por múltiples factores (atencionales, perceptivos, de lenguaje o de pensamiento) y que en el mejor de los casos solo constituye una mala caricatura de la realidad vivida por el testigo. Véase Manzanero (2010a).

lógica, el sentido común y la experiencia para determinar la credibilidad, la veracidad y el peso probatorio de las pruebas presentadas (Abel Lluch, 2020). Para realizar una valoración racional de la prueba testifical se deben considerar diversos factores, como la coherencia interna de las pruebas, la consistencia con otros elementos probatorios, la congruencia con la lógica y la experiencia común, la credibilidad de los testigos o peritos, y cualquier otro elemento que pueda influir en la fiabilidad de la prueba. El objetivo es llegar a una conclusión fundamentada que permita acercarse lo más posible a la verdad, teniendo en cuenta el conjunto de pruebas presentadas en el caso y aplicando los estándares legales correspondientes. Sin embargo, como cualquier valoración racional que requerirá de la utilización de generalizaciones, de máximas de experiencia (Vázquez, 2015), cabe preguntarse si debemos confiar en nuestra intuición para tomar decisiones sobre la fiabilidad de la prueba testifical o si las máximas de la experiencia que intuitivamente se asumen por los operadores del sistema de justicia para valorar la prueba testifical tienen su debido sustento científico. Wells y Lindsay (1983)³ propusieron el *modelo del enjuiciador de memoria*, que permite entender el tipo de razonamiento intuitivo utilizado para valorar la prueba testifical. Según este modelo, el principal indicador de credibilidad sería la información condicional, que responde a la pregunta: «¿Yo qué recordaría si fuese ese testigo y hubiera vivido esa experiencia?». Se trata de la aplicación de la empatía a la valoración de la prueba testifical. El problema es que esa operación se podría realizar desde los estereotipos sobre cómo funciona la memoria de los testigos más que desde las evidencias científicas. Esta propuesta ha dado lugar y justifica los estudios que tienen como objetivo valorar las creencias sobre el funcionamiento de la memoria de los testigos.

La ciencia en colaboración con el derecho, en lo que se ha venido a denominar la psicología del testimonio (Manzanero, 2008), que aplica los conocimientos sobre el funcionamiento normal y patológico de los procesos cognitivos al desarrollo de procedimientos específicos para optimizar la obtención y valoración de los testimonios (González y Manzanero, 2018; Manzanero y González, 2015) con el objetivo principal de obtener la mayor cantidad de información maximizando su exactitud. Dos son, por lo tanto, los objetos de estudio de esta disciplina: la exactitud y la credibilidad de declaraciones e identificaciones. Esta especialidad de la psicología (Muñoz *et al.*, 2011) tiene su origen en el siglo xix de la mano de los avances en la psicología científica respecto del estudio del funcionamiento de la percepción y la memoria humanas (Manzanero, 2010b). Desde entonces se han establecido como evidencias empíricas algunos fenómenos, entre los que se pueden citar el efecto del paso del tiempo o de la información falsa postsuceso en la exactitud de las testificales. Profesionales bien informados acerca de la psicología del testimonio conocerán estas evidencias. No obstante, algunos operadores jurídicos podrían desconocerlas, afectando a la obtención y valoración de la prueba testifical.

³ Aunque es de hace ya unos años, sigue tratándose del modelo que de forma más plausible y sencilla intenta explicar los procesos intuitivos en la valoración de los recuerdos aportados por terceras personas y, por lo tanto, el más aplicable a la valoración de los testimonios.

En este marco, diferentes estudios (Benton *et al.*, 2006; Houston *et al.*, 2013; Kassir *et al.*, 2001; Wells, 1984; Wise y Safer, 2004; Yarmey y Jones, 1983) han explorado las teorías intuitivas que los profesionales vinculados con el sistema de justicia tienen acerca de la memoria de los testigos. En general, los resultados muestran que existen falsas creencias contrarias a la evidencia científica. Sin embargo, lo más sorprendente es que estas creencias son compartidas por aquellas personas que, debido a su profesión, están en contacto directo con testigos (forenses, policías, abogados, fiscales y jueces). Wells (1984), en un estudio realizado en Canadá, concluyó que la intuición humana no es suficiente para juzgar la memoria de los testigos, ya que la mayoría de los encuestados no comprende la sobreestimación de la duración de eventos por parte de los testigos, la influencia de la exposición a fotografías antes de una rueda de identificación, la falta de superioridad de la policía cuando sus agentes actúan como testigos, el impacto de la gravedad del delito, las distorsiones producidas por la hipnosis y la ineficacia de esta, ni el efecto de las instrucciones sesgadas en una rueda de reconocimiento. Kassir *et al.* (2001) analizaron las creencias sobre aspectos relativos a los factores que influyen en la exactitud de la prueba testifical (declaraciones e identificaciones) mediante un cuestionario que contestaron sesenta y cuatro psicólogos relacionados con el sistema de justicia. El estudio mostró un ligero avance en el conocimiento respecto a una encuesta previa (Kassir *et al.*, 1989) y consenso respecto a muchos fenómenos relacionados con los testimonios de testigos presenciales. Granhag *et al.* (2005) examinaron las creencias de profesionales del derecho suecos sobre el testimonio de testigos oculares, mostrando escepticismo, especialmente entre jueces, sobre la fiabilidad de los testimonios de niños. Benton *et al.* (2006) evaluaron si las creencias sobre la memoria de los testigos habían cambiado con el tiempo entre jueces, jurados y abogados. Desafortunadamente, los resultados fueron desalentadores, con un 84% de jurados y un 60% de jueces y abogados en desacuerdo con los expertos. En una comparación entre fiscales y abogados defensores (Wise *et al.*, 2009), se encontró que los fiscales estaban menos informados sobre el funcionamiento de la memoria de los testigos y menos dispuestos a permitir el testimonio de expertos en este campo en comparación con los abogados defensores. Magnussen *et al.* (2010), en Noruega, encontraron que tanto jueces como jurados tenían conocimientos limitados sobre psicología del testimonio. La experiencia previa como jurado no se correlacionó con el conocimiento. Estos resultados resaltan la necesidad de mayor formación y conciencia sobre la memoria de los testigos en el sistema judicial.

Los tópicos explorados en estos estudios en general tratan sobre las diligencias de reconocimiento de sospechosos, los factores que influyen en la exactitud de los testimonios y su valoración. Sobre esto, la psicología del testimonio y la psicología experimental (que abarca el estudio sobre el funcionamiento normal y patológico de procesos cognitivos como la atención, la percepción y la memoria, y el papel que la motivación, la emoción, el lenguaje o los procesos de pensamiento tienen sobre ellos) ha establecido un corpus afianzado de conocimiento. Entre otras evidencias científicas, conviene recordar que la percepción y la memoria son procesos constructivos y dinámicos, que no reflejan la realidad sino una interpretación sesgada de

la misma, que las características de los testigos (edad, conocimientos y experiencias previas, estado mental...), las características del suceso (frecuencia, grado de violencia, duración...), los procedimientos de toma de declaración (preguntas cerradas y sugestivas, reiteración...) y otros factores como el paso del tiempo pueden afectar negativamente a la exactitud de los testimonios. Entre otras, la ciencia ha establecido la existencia de falsas memorias, la falta de capacidad para generar recuerdos hasta los seis años de edad, o no ha podido encontrar evidencias sobre fenómenos como las memorias reprimidas. De igual modo, ha demostrado la ineficacia de la hipnosis como método para obtener testimonios y que los errores al identificar visualmente a un desconocido son muy frecuentes, al tiempo que ha sido incapaz de desarrollar adecuados procedimientos para la detección de la mentira⁴.

En el presente estudio utilizamos el cuestionario de Kassir *et al.* (2001) para evaluar las creencias de policías (con y sin experiencia), jueces en formación en la Escuela Judicial, dependiente del Consejo General del Poder Judicial, y estudiantes de grados relacionados con las ciencias sociales y jurídicas (Derecho, Criminología y Psicología). Como en los estudios sobre creencias realizados en décadas anteriores, esperamos encontrar importantes discrepancias con las evidencias científicas que no se verán influidas por el tipo de actividad de los participantes. El objetivo principal será determinar las creencias que distintos operadores jurídicos podrían utilizar para valorar la prueba testifical y que, si no coinciden con las evidencias científicas, podrían llevar a conclusiones erróneas.

2. MÉTODO

2.1. Participantes

Participaron en el estudio 695 personas⁵, miembros de las Fuerzas y Cuerpos de Seguridad del Estado (Policía Nacional y Guardia Civil), jueces en formación de la Escuela Judicial y estudiantes de grado del área de las ciencias sociales y jurídicas (Criminología, Derecho y Psicología) de las universidades Complutense de Madrid, Salamanca y La Laguna. Un 35.8% de los participantes eran hombres, un 63.9% mujeres y un 0.3% marcaron la opción de «otro». De los 213 participantes de CNP y GC, 120 eran funcionarios con experiencia ($M = 15.26$ años, $DT = 7.91$) y 93 estaban realizando la formación básica en las respectivas academias después de aprobar las oposiciones. La participación fue voluntaria y anónima.

⁴ Para profundizar en lo que la investigación científica desde la psicología del testimonio ha aportado, puede consultarse Manzanero (2010a).

⁵ Para una población aproximada de ochenta mil personas (egresados anuales en ciencias sociales y jurídicas, y alumnos de las academias de policía y de la Escuela Judicial), con un nivel de confianza del 95% y un margen de error de 5%, la muestra mínima necesaria sería de 397 participantes (Arkin y Colton, 1962).

TABLA 1.
Tamaños de los grupos muestrales (N), sus edades medias (M) y desviaciones típicas (DT).

Grupo	Edad		N
	M	DT	
Escuela Judicial	28.92	3.20	99
Policías con experiencia	40.13	7.57	120
Academia de policía	28.56	4.93	93
Criminología	21.97	5.91	82
Derecho	20.48	6.06	106
Psicología	21.52	6.15	195
Total	26.62	9.12	695

2.2. Procedimiento

Se utilizó una traducción al castellano del cuestionario en torno a las creencias sobre el funcionamiento de la memoria de los testigos elaborado por Kassin *et al.* (2001), que consta de treinta preguntas. En el Anexo pueden verse las respuestas consideradas correctas en función de las evidencias científicas disponibles.

Las preguntas formuladas aluden a aspectos relacionados con las diligencias de reconocimiento y su valoración (2, 3, 4, 5, 8, 12, 18, 19, 21, 27, 28 y 30), a los factores que influyen en la exactitud de los testimonios (1, 6, 10, 13, 16, 17, 20 y 24), al funcionamiento de la memoria (7, 9, 22), a la capacidad de los testigos para testificar en función de la edad (23, 25, 26 y 29), y a la utilidad de determinados procedimientos para obtener los testimonios (11, 14 y 15).

Las encuestas se aplicaron de forma individual mediante la herramienta *google forms*, a la que los participantes accedieron a través de un código QR. No se establecieron tiempos límite para su contestación, que en ningún caso excedió los diez minutos.

3. RESULTADOS

3.1. Aciertos generales en el cuestionario

La media de aciertos globales fue de 17.32 ($DT = 3.56$). Teniendo en cuenta que el cuestionario incluía treinta preguntas, el porcentaje medio de aciertos globales fue de solo un 57.74%.

El análisis de cada una de las preguntas formuladas (ver Tabla 2) indica que en general los participantes contestaron correctamente (cuartil Q1) al efecto de la intoxicación alcohólica, el estrés, la redacción de preguntas, las actitudes y expectativas, la información posterior al evento, el sesgo inducido por fotografías policiales, la sugestionabilidad infantil y las instrucciones para la formación de ruedas.

Los errores se incrementaron (cuartil Q2) respecto al efecto del enfoque en el arma, la maleabilidad de la confianza, los falsos recuerdos de la infancia, el tiempo de exposición, la discriminabilidad entre los testimonios verdaderos y falsos, la confusión del autor con otros testigos, los testigos de edad avanzada, las ruedas de fotos, y el sesgo racial cruzado.

El porcentaje de respuestas correctas no alcanzó el 50% (cuartil Q3) sobre el efecto del paso del tiempo, de la violencia, la relación exactitud-confianza, la formación de ruedas con descripciones coincidentes, la precisión del testimonio infantil, la sugestionabilidad de la hipnosis, la validez del tiempo de respuesta en las identificaciones como indicador de exactitud, el formato de presentación de las ruedas, la exactitud de la hipnosis, la equidad en la formación de ruedas y el entrenamiento de los testigos.

Finalmente, los participantes tenían creencias mayoritariamente contrarias a la ciencia (cuartil Q4) respecto a la percepción del color y las memorias reprimidas.

TABLA 2.
Porcentajes y cantidad de aciertos generales para cada una de las preguntas formuladas, ordenadas de mayor a menor.

Cuartil	Pregunta	Aciertos	%
Q1	20. Intoxicación alcohólica	679	97.70
	1. Estrés	670	96.40
	11. Redacción de preguntas	668	96.12
	16. Actitudes y expectativas	642	92.37
	9. Información posterior al evento	632	90.94
	21. Sesgo inducido por fotografías policiales	573	82.45

Cuartil	Pregunta	Aciertos	%
Q1	26. Sugestionabilidad infantil	551	79.28
	5. Instrucciones para la formación de ruedas	522	75.11
Q2	2. Enfoque en el arma	520	74.82
	19. Maleabilidad de la confianza	497	71.51
	23. Falsos recuerdos de la infancia	466	67.05
	6. Tiempo de exposición	457	65.76
	24. Discriminabilidad verdad-falsedad	456	65.61
	12. Confusión del autor con otros testigos	419	60.29
	3. Rueda de fotos	402	57.84
	29. Testigos de edad avanzada	389	55.97
	18. Sesgo racial cruzado	352	50.65
Q3	7. Curva de olvido	346	49.78
	17. Efecto violencia	342	49.21
	27. Formación de ruedas con descripciones coincidentes	306	44.03
	8. Relación exactitud-confianza	305	43.88
	25. Precisión del testimonio de un niño	276	39.71
	15. Sugestionabilidad de la hipnosis	267	38.42
	30. Velocidad de identificación	244	35.11
	28. Formato de presentación de ruedas	242	34.82
	14. Exactitud de la hipnosis	240	34.53
	4. Equidad en la formación de ruedas	207	29.78
	13. Testigos entrenados	200	28.78
Q4	10. Percepción del color	108	15.54
	22. Memorias reprimidas	51	7.34

3.2. Policías vs. estudiantes de grado vs. jueces en formación

Considerando los seis grupos, se encontraron efectos del tipo de actividad, $F(5,689) = 5.846$, $p < .001$, $\eta^2 = .041$. El análisis de los aciertos medios obtenidos por cada uno de los grupos de participantes muestra que los estudiantes de Crimino-

logía y los jueces en formación de la Escuela Judicial acertaron más que los estudiantes de Derecho, (Scheffe = 1.998, $p < .01$; Scheffe = 1.725, $p < .05$) y de Psicología (Scheffe = 1.854, $p < .01$; Scheffe = 1.581, $p < .05$). ¿No deParece que una formación mixta específica en Psicología, Derecho y Ciencias Policiales podría proporcionar un mejor conocimiento sobre el funcionamiento de la memoria de los testigos (véase Tabla 3). No se encontraron diferencias entre los policías con experiencia y los de la academia con el resto de los grupos.

TABLA 3.
Porcentajes (%) de aciertos, puntuaciones medias (M) y sus respectivas desviaciones típicas (DT) en función del tipo de actividad.

	%	DT	M	DT	N
Escuela Judicial	60.87	11.65	18.26	3.49	99
Policía con experiencia	58.39	12.45	17.52	3.73	120
Academia policía	57.45	11.08	17.25	3.32	93
Criminología	61.79	11.90	18.54	3.57	82
Derecho	55.13	11.53	16.54	3.46	106
Psicología	55.61	11.47	16.68	3.44	195
Total	57.74		17.32	3.56	695

En la Tabla 4 (ver página siguiente) pueden observarse los porcentajes de aciertos para cada una de las preguntas en función de la actividad de los participantes. Se encontraron efectos del tipo de actividad de los participantes para veintidós de las treinta preguntas formuladas.

3.3. Jueces en formación

Los jueces en formación de la Escuela Judicial contestaron por debajo del 50% de aciertos (Q3) las preguntas sobre la exactitud de la hipnosis, los testigos de edad avanzada, la sugestionabilidad de la hipnosis, la formación de ruedas con descripciones coincidentes, la relación exactitud-confianza, la precisión del testimonio infantil, el formato de presentación de ruedas, la equidad en la formación de ruedas, la validez del tiempo de respuesta en las identificaciones como indicador de exactitud y los testigos entrenados.

Las preguntas sobre la percepción del color y las memorias reprimidas fueron contestadas erróneamente por casi la totalidad de los jueces en formación (Q4).

TABLA 4.
Porcentaje de aciertos para cada pregunta en función de la actividad de los participantes.

N = 211	Escuela Judicial (n = 99)					Academia (n = 93)					Policía (n = 120)					Criminología (n = 82)					Derecho (n = 106)					Psicología (n = 195)					X ²	p
	1. Estrés *	2. Enfoque en el arma **	3. Rueda de fotos **	4. Equidad en la formación de ruedas **	5. Instrucciones para la formación de ruedas **	6. Tiempo de exposición **	7. Curva de olvido	8. Relación exactitud-confianza **	9. Información posterior al evento **	10. Percepción del color	11. Redacción de preguntas **	12. Confusión del autor con otros testigos **	13. Testigos entrenados **	14. Exactitud de la hipnosis **	15. Sugestionabilidad de la hipnosis **	16. Actitudes y expectativas *	17. Efecto violencia *															
	97.98	89.90	68.69	35.35	81.82	64.65	41.41	92.93	18.18	98.99	66.67	29.29	47.47	45.45	92.93	59.60																
	98.92	82.80	66.67	47.31	73.12	80.65	36.56	91.40	13.98	98.92	50.54	18.28	12.90	25.81	94.62	58.06																
	98.33	75.00	64.17	45.00	79.17	80.00	48.33	98.33	11.67	98.33	55.83	20.00	15.83	22.50	92.50	45.00																
	98.78	95.12	62.20	18.29	86.59	67.07	58.54	96.34	8.54	98.78	79.27	40.24	41.46	56.10	96.34	56.10																
	93.40	58.49	50.94	28.30	54.72	57.55	33.02	81.13	18.87	84.91	58.49	35.85	43.40	40.57	83.96	44.34																
	93.85	63.59	46.15	14.87	76.41	54.36	45.64	88.21	18.46	96.92	57.44	30.26	42.05	42.05	93.85	42.05																
	11.481	61.087	23.344	54.439	33.167	34.506	15.711	25.498	7.298	43.310	19.514	17.579	55.442	33.305	13.794	14.602																
	.043	.000	.000	.000	.000	.000	.008	.000	.199	.000	.002	.004	.000	.000	.017	.012																

18. Sesgo racial cruzado **	55.56	52.69	67.50	60.98	34.91	41.03	35.974	.000
19. Maleabilidad de la confianza	74.75	70.97	73.33	74.39	63.21	72.31	4.700	.454
20. Intoxicación alcohólica	97.98	98.92	99.17	98.78	94.34	97.44	7.610	.179
21. Sesgo inducido por fotografías policiales	80.81	89.25	87.50	78.05	78.30	81.03	7.899	.162
22. Memorias reprimidas **	2.02	4.30	5.83	7.32	15.09	8.21	15.372	.009
23. Falsos recuerdos de la infancia **	70.71	41.94	56.67	76.83	77.36	73.85	45.731	.000
24. Discriminabilidad verdad-falsedad **	68.69	50.54	44.17	80.49	70.75	75.38	51.780	.000
25. Precisión del testimonio infantil *	40.40	45.16	46.67	43.90	41.51	29.74	12.435	.029
26. Sugestionabilidad infantil	84.85	81.72	80.00	84.15	73.58	75.90	6.877	.230
27. Formación de ruedas con descripciones coincidentes **	45.45	61.29	55.00	30.49	47.17	32.31	34.584	.000
28. Formato de presentación de ruedas	40.40	32.26	36.67	31.71	39.62	30.77	4.647	.461
29. Testigos de edad avanzada	47.47	55.91	58.33	62.20	58.49	54.87	4.830	.437
30. Velocidad de identificación *	33.33	40.86	46.67	26.83	33.96	30.26	13.067	.023

Efectos estadísticamente significativos * $p < .05$; ** $p < .01$

3.4. Policías en función de la experiencia

No se encontraron efectos estadísticamente significativos en función de la experiencia de los policías. Los policías con experiencia y los que aprobaron recientemente las oposiciones y están en la academia en el periodo de formación acertaron de media la misma cantidad de preguntas del cuestionario; $F(1,211) = 0.324$, $p = .570$, $\eta^2 = .002$.

Solo se encontraron diferencias significativas para el efecto de la información posterior al evento ($\chi^2 = 5.633$, $p < .05$), el sesgo racial cruzado ($\chi^2 = 4.833$, $p < .05$) y los falsos recuerdos en la infancia ($\chi^2 = 4.548$, $p < .05$). En los tres casos, los policías con experiencia acertaron más que los policías sin experiencia (Tabla 4).

Las preguntas sobre la discriminabilidad verdad-falsedad, la equidad en la formación de ruedas, la curva de olvido, la precisión del testimonio de un niño, los falsos recuerdos de la infancia, la validez del tiempo de respuesta en las identificaciones como indicador de exactitud, la relación exactitud-confianza, y el formato de presentación de ruedas fueron contestadas erróneamente por más de la mitad de los policías participantes (Q3).

Los peores resultados (Q4) se encontraron respecto a la sugestionabilidad de la hipnosis, los testigos entrenados, la percepción del color, la exactitud de la hipnosis y las memorias reprimidas.

3.5. Estudiantes en función de la especialidad

Se encontraron efectos estadísticamente significativos del tipo de estudio sobre la cantidad de aciertos totales en el cuestionario. Como puede verse en la Tabla 7, los estudiantes del Grado en Criminología acertaron de media más que los estudiantes del Grado en Derecho y Psicología; $F(2, 380) = 9.754$, $p < .001$, $\eta^2 = .049$.

En la Tabla 8 pueden apreciarse las diferencias para cada pregunta en función del tipo de estudio. Los estudiantes de Criminología acertaron más en las preguntas sobre el efecto del arma ($\chi^2 = 34.257$, $p < .01$), las instrucciones de identificación ($\chi^2 = 26.518$, $p < .01$), la relación confianza-exactitud en los reconocimientos ($\chi^2 = 12.254$, $p < .01$), el efecto de la información postsuceso ($\chi^2 = 10.138$, $p < .01$), la confusión del autor con otros testigos ($\chi^2 = 12.683$, $p < .01$) y el sesgo racial cruzado ($\chi^2 = 13.851$, $p < .01$); y acertaron menos sobre el efecto del paso del tiempo ($\chi^2 = 8.138$, $p < .05$). Los estudiantes de Derecho acertaron más en la pregunta sobre la equidad en la formación de las ruedas de reconocimiento ($\chi^2 = 8.017$, $p < .05$), la formación de ruedas con descripciones coincidentes ($\chi^2 = 7.972$, $p < .05$), y menos sobre el efecto de la redacción de las preguntas ($\chi^2 = 21.803$, $p < .01$) y el efecto de las actitudes y expectativas ($\chi^2 = 11.767$, $p < .01$). Los estudiantes de Psicología acertaron menos sobre la precisión del testimonio infantil ($\chi^2 = 6.929$, $p < .05$).

Más de la mitad de los estudiantes (Q3) en general contestaron erróneamente las preguntas sobre la curva de olvido, el efecto violencia, la sugestionabilidad de la hipnosis, la relación exactitud-confianza, el sesgo racial cruzado, la exactitud de la hipnosis, la precisión del testimonio de un niño, la formación de ruedas con descripciones coincidentes, los testigos entrenados, el formato de presentación de ruedas y la validez del tiempo de respuesta en las identificaciones como indicador de exactitud.

Casi todos los estudiantes (Q4) contestaron erróneamente las preguntas sobre la equidad en la formación de ruedas, la percepción del color y las memorias reprimidas.

4. DISCUSIÓN

Los resultados del presente estudio han mostrado que son frecuentes las falsas creencias sobre el funcionamiento de la memoria de los testigos, los procedimientos para obtener y valorar los testimonios y los factores que influyen en la exactitud de declaraciones e identificaciones. Solo un poco más de la mitad de las preguntas formuladas fue correctamente contestada por encima del azar. Trece de las treinta preguntas fueron erróneamente contestadas por más de la mitad de los participantes. La mayoría de los errores estaban relacionados con el funcionamiento de la memoria, con un porcentaje medio de aciertos del 49.35%, y las diligencias de identificación, con un porcentaje medio del 50.02%. La utilidad de determinados procedimientos para obtener los testimonios fue correctamente contestada por el 56.35% de los participantes. Sobre la capacidad de los testigos para testificar en función de la edad se encontró un conocimiento desigual. El conocimiento sobre la precisión de los testimonios infantiles alcanzó solo el 39.71%, aunque cerca del 80% contestó adecuadamente a la pregunta sobre la sugestionabilidad infantil. El 45% mostró un conocimiento inadecuado sobre la precisión de los testimonios en personas mayores en comparación con los aportados por personas más jóvenes. Las respuestas correctas sobre los factores que influyen en la exactitud de los testimonios alcanzaron el 63.92%. Las preguntas sobre las memorias reprimidas y la adaptación al color fueron erróneamente contestadas por casi todos los participantes independientemente del colectivo al que pertenecían.

En general, las creencias sobre estas cuestiones fueron compartidas por los distintos colectivos participantes y no se advirtieron diferencias reseñables. El mejor rendimiento se encontró en los jueces en formación y los estudiantes de Criminología, que reciben formación combinada de Derecho y Psicología. No se encontraron diferencias con los grupos de policías que obtuvieron un rendimiento a mitad de camino entre aquellos y los estudiantes de Derecho y Psicología. Estos resultados son congruentes con los encontrados en estudios previos realizados con profesionales de distintos países a lo largo de las últimas décadas (Benton *et al.*, 2006; Houston *et al.*,

2013; Kassin *et al.*, 2001; Read y Desmarais, 2009; Wells, 1984; Wise y Safer, 2004; Yarmey y Jones, 1983).

Los datos muestran la necesidad de especialización también entre los psicólogos, de los que se esperaría un mayor conocimiento sobre la memoria de los testigos y son congruentes con los estudios previos que muestran discrepancias sobre importantes tópicos sobre el funcionamiento de la memoria⁶ y los mejores procedimientos para la obtención y valoración de los testimonios más allá de la experiencia (González y Manzanero, 2018; Scott *et al.*, 2014).

De acuerdo con el *Modelo de Enjuiciador de Memoria* (Wells y Lindsay, 1983) se alcanzarían conclusiones erróneas en la valoración intuitiva de los testimonios en tanto se realice a partir de creencias contrarias a las evidencias científicas. Un reciente estudio (Schell-Leugers *et al.*, 2023) sobre las creencias y prácticas de toma de declaración en la policía española mostró que creen ser capaces de detectar la mentira, al tiempo que muestran falsas creencias a este respecto. Por esta razón se hace necesaria una valoración experta de los testimonios que deberían llevar a cabo profesionales especialistas en psicología del testimonio, tal y como ya recoge el Anteproyecto de Reforma de la LECrim (2020).

En este sentido, resulta de especial interés la distinción entre ciencia y pseudo-ciencia (Álvarez, 2018; Bunge, 2014; Munro y Munro, 2014), ya que esta última puede ser la responsable de muchas de las falsas creencias sobre el funcionamiento de la memoria de los testigos.

Por esta razón, los resultados apuntan a la necesidad de la formación especializada que combine psicología y derecho. Los pobres resultados de los estudiantes de Derecho, excepto acerca de cómo realizar una rueda de reconocimiento, muestran que la formación de los juristas debería incluir en sus planes de estudio asignaturas de psicología del testimonio. El pobre rendimiento de los estudiantes en Psicología, excepto sobre el efecto del paso del tiempo o el tipo de preguntas sobre la memoria, muestra que la formación de los psicólogos que se dediquen al ámbito forense debería incluir formación especializada (Muñoz *et al.*, 2011), y que la formación general o en otras especialidades (clínica, social, educativa...) no es suficiente. De igual forma, se pone de manifiesto la necesidad de mejorar los conocimientos que los miembros de las Fuerzas y Cuerpos de Seguridad del Estado tienen sobre esta materia, tal y como han reconocido ellos mismos en estudios previos (Schell-Leugers *et al.*, 2023).

En cualquier caso, cabe destacar que las principales limitaciones de este tipo de trabajo se encuentran en la formulación de las preguntas incluidas en la encuesta, así como en los participantes. Read y Desmarais (2009) analizaron si introducir

⁶ Véase, por ejemplo, la polémica sobre las memorias reprimidas en Brand, Schielke y Brams (2017); Brand, Schielke, Brams y Dicomio (2017); Manzanero y Morales-Valiente (2024); Manzanero y Palomo (2020); Merckelbach y Patihis (2018); Patihis *et al.* (2021); Patihis *et al.* (2019).

Véase, también, la polémica sobre la detección de la mentira en Denault *et al.* (2020); Masip *et al.* (2011); Sánchez y Manzanero (2023).

cambios en la formulación de las preguntas y si permitir distintos tipos de respuesta podría afectar a los resultados al evaluar a participantes legos. Los resultados mostraron algunos efectos, mejorando el rendimiento, pero los resultados principales se mantuvieron. Así, los participantes demostraron tener escasos conocimientos sobre aproximadamente el 50% de los tópicos evaluados.

Futuras investigaciones podrían incluir a profesionales con experiencia en el ámbito de la justicia a fin de analizar en qué medida la experiencia afecta al conocimiento sobre el funcionamiento de la memoria de los testigos.

5. CONCLUSIONES

Los resultados del estudio señalan la prevalencia de falsas creencias sobre el funcionamiento de la memoria de los testigos, los factores que afectan la precisión de las declaraciones e identificaciones y los procedimientos para obtener y valorar los testimonios. Cerca del 50% de las preguntas fueron respondidas incorrectamente. Las áreas con mayor proporción de errores estaban relacionadas con el funcionamiento de la memoria y los procedimientos de identificación. Los jueces en formación y los estudiantes de Criminología mostraron el mejor rendimiento, mientras que los policías se ubicaron a medio camino entre estos y los estudiantes de Derecho y Psicología. En cualquier caso, no hubo grandes diferencias entre los diversos colectivos participantes en cuanto a creencias erróneas, lo que sugiere que muchas son compartidas de manera generalizada. La conclusión es coherente con estudios previos realizados en distintos países a lo largo de las últimas décadas. La probabilidad de llegar a conclusiones erróneas en la valoración intuitiva de testimonios será alta si se basa en creencias contrarias a las evidencias científicas, por lo que se hará necesaria una valoración experta de los testimonios a cargo de profesionales especialistas en psicología del testimonio, tal como propone el Anteproyecto de Reforma de la LECrim (2020). En cualquier caso, destaca la necesidad de una formación especializada que combine psicología y derecho, especialmente en el caso de los juristas, y la importancia de una formación más específica para psicólogos que se dediquen al ámbito forense, y no solo formación general o en otras especialidades. En resumen, los resultados indican la urgencia de mejorar la formación en psicología del testimonio en diversas disciplinas para abordar las creencias erróneas y mejorar la calidad de la valoración de testimonios en el sistema judicial.

BIBLIOGRAFÍA

- Abel Lluch, X. (2020). *La valoración de la credibilidad del testimonio*. Wolters Kluwer.
- Álvarez, M. A. (2018). *Datos blandos para ciencias duras. Psicología y neurociencias*. Editorial EOS.
- Arkin, H. y Colton, R. R. (1962). *Tables for Statisticians*. Barnes & Noble.

- Benton, T. R., Ross, D. F., Bradshaw, E., Thomas, W. N. y Bradshaw, G. S. (2006). Eyewitness Memory Is Still Not Common Sense: Comparing Jurors, Judges and Law Enforcement to Eyewitness Experts. *Applied Cognitive Psychology*, 20, 115-129. <https://doi.org/10.1002/acp.1171>
- Brand, B. L., Schielke, H. J. y Brams, J. S. (2017). Assisting the Courts in Understanding and Connecting with Experiences of Disconnection: Addressing Trauma-Related Dissociation as a Forensic Psychologist, part I. *Psychological Injury and Law*, 10, 283-297. <https://doi.org/10.1007/s12207-017-9304-8>
- Brand, B. L., Schielke, H. J., Brams, J. S. y Dicomio, R. A. (2017). Assessing Trauma-Related Dissociation in Forensic Contexts: Addressing Trauma-Related Dissociation as a Forensic Psychologist, part II. *Psychological Injury and Law*, 10, 298-312. <https://doi.org/10.1007/s12207-017-9305-7>
- Bunge, M. (2014). *Las pseudociencias. ¡Vaya timo!*. Laetoli.
- De Paula Ramos, V. (2019). *La prueba testifical: del subjetivismo al objetivismo, del aislamiento científico al diálogo con la psicología y la epistemología*. Marcial Pons.
- Denault, V., Plusquellec, P., Jupe, L. M., St-Yves, M., Dunbar, N. E., (...), Hartwig, M., Sporer, S. L., Rioux-Turcotte, J., Jarry, J., Walsh, D., Otgaar, H., Viziteu, A., Talwar, V., Keatley, D. A., Blandón-Gitlin, I., Townson, C., Deslauriers-Varin, N., Lilienfeld, S. O., Patterson, M. L., Areh, I. et al. (2020). The Analysis of Nonverbal Communication: The Dangers of Pseudoscience in Security and Justice Contexts. *Anuario de Psicología Jurídica*, 30, 1-12. <https://doi.org/10.5093/apj2019a9>
- González, J. L. y Manzanero, A. L. (2018). *Obtención y valoración del testimonio. Protocolo Holístico de Evaluación de la Prueba Testifical (HELPT)*. Pirámide.
- Granhag, P. A., Strömwall, L. A. y Hartwig, M. (2005). Eyewitness Testimony: Tracing the Beliefs of Swedish Legal Professionals. *Behavioral Sciences & the Law*, 23(5), 709-727. <https://doi.org/10.1002/bsl.670>
- Houston, K. A., Hope, L., Memon, A. y Read, J. D. (2013). Expert Testimony on Eyewitness Evidence: In Search of Common Sense. *Behavioral sciences & the law*, 31(5), 637-651. <https://doi.org/10.1002/bsl.2080>
- Kassin, S. M., Elisworth, P. C. y Smith, V. L. (1989). The «General Acceptance» of Psychological Research on Eyewitness Testimony: A Survey of the Experts. *American Psychologist*, 44, 1089-1098. <https://psycnet.apa.org/doi/10.1037/0003-066X.44.8.1089>
- Kassin, S. M., Tubbs, V. A., Hosch, H. M. y Memon, A. (2001). On the «General Acceptance» of Eyewitness Testimony Research: A New Survey of the Experts. *American Psychologist*, 56(5), 405-416. <https://doi.org/10.1037/0003-066X.56.5.405>
- Magnussen, S., Melinger, A., Stridbeck, U. y Raja, A. Q. (2010). Beliefs about Factors Affecting the Reliability of Eyewitness Testimony: A Comparison of Judges, Jurors and the General Public. *Applied Cognitive Psychology*, 24(1), 122-133. <https://doi.org/10.1002/acp.1550>
- Manzanero, A. L. (2008). *Psicología del Testimonio. Una aplicación de los estudios sobre la memoria*. Pirámide.
- Manzanero, A. L. (2010a). *Memoria de testigos. Obtención y valoración de la prueba testifical*. Pirámide.
- Manzanero, A. L. (2010b). Hitos de la historia de la psicología del testimonio en la escena internacional. *Boletín de Psicología*, 100, 89-104. <https://docta.ucm.es/bitstreams/7d89ad17-663c-4301-8828-8031554a3050/download>
- Manzanero, A. L. y González, J. L. (2015). Modelo holístico de evaluación de la prueba testifical (HELPT). *Papeles del Psicólogo*, 36(2), 125-138. <https://www.papelesdelpsicologo.es/pdf/2568.pdf>
- Manzanero, A. L. y Morales-Valiente, C. (2024). Memory Wars: A Solution to the Conflict. *Papeles del Psicólogo/ Psychologist Papers*, 45(1), 34-38. <https://doi.org/10.23923/pap.psicol.3030>
- Manzanero, A. L. y Palomo, R. (2020). Dissociative Amnesia: Beyond the Evidence about the Functioning of Memory. *Anuario de Psicología Jurídica*, 30, 43-46. <https://doi.org/10.5093/apj2019a14>
- Masip, J., Herrero, C., Garrido, E. y Barba, A. (2011). Is the Behavior Analysis Interview Just Common Sense? *Applied Cognitive Psychology*, 25, 593-604. <https://doi.org/10.1002/acp.1728>
- Merckelbach, H. y Patihis, L. (2018). Why «Trauma-Related Dissociation» is a Misnomer in Courts: A Critical Analysis of Brand et al. (2017). *Psychological Injury and Law*, 11, 370-376. <https://doi.org/10.1007/s12207-018-9328-8>

- Munro, G. D. y Munro, C. A. (2014). «Soft» versus «hard» Psychological Science: Biased evaluations of scientific evidence that threatens or supports a strongly held political identity. *Basic and Applied Social Psychology*, 36(6), 533-543. <https://doi.org/10.1080/01973533.2014.960080>
- Muñoz, J. M., Manzanero, A. L., Alcázar, M. A., González, J. L., Pérez, M. L. y Yela, M. (2011). Psicología jurídica en España: delimitación conceptual, campos de investigación e intervención y propuesta formativa de la enseñanza oficial. *Anuario de Psicología Jurídica*, 21, 3-14. <https://doi.org/10.5093/jr2011v21a1>
- Nieva Fenoll, J. (2010). *La valoración de la prueba*. Marcial Pons.
- Patihis, L., Ho, L. Y., Loftus, E. F. y Herrera, M. E. (2021). Memory Experts' Beliefs about Repressed Memory. *Memory*, 29(6), 823-828. <https://doi.org/10.1080/09658211.2018.1532521>
- Patihis, L., Otgaar, H. y Merckelbach, H. (2019). Expert Witnesses, Dissociative Amnesia, and Extraordinary Remembering: Response to Brand *et al.* *Psychological Injury and Law*, 12, 281-285. <https://doi.org/10.1007/s12207-019-09348-8>
- Ramírez Ortiz, J. L. (2020). El testimonio único de la víctima en el proceso penal desde la perspectiva de género. *Quaestio Facti. Revista Internacional Sobre Razonamiento Probatorio*, 2020(1), 201-246. https://doi.org/10.33115/udg_bib/qf.i0.22288
- Read, J. D. y Desmarais, S. L. (2009). Lay Knowledge of Eyewitness Issues: A Canadian Evaluation. *Applied Cognitive Psychology*, 23(3), 301-326. <https://doi.org/10.1002/acp.1459>
- Sánchez, N. y Manzanero, A. L. (2023). El engaño en contextos judiciales. *Revista Ítalo-Española de Derecho Procesal*, 2023(1), 5-22. <https://doi.org/10.37417/rivitsproc/1518>
- Schell-Leugers, J. M., Masip, J., González, J. L., Vanderhallen, M. y Kassin, S. M. (2023). Police Interviewing in Spain: A Self-report Survey of Police Practices and Beliefs. *Anuario de Psicología Jurídica*, 33(1), 27-40. <https://doi.org/10.5093/apj2022a4>
- Scott, M. T., Manzanero, A. L., Muñoz, J. M. y Köhnken, G. (2014). Admisibilidad en contextos forenses de indicadores clínicos para la detección del abuso sexual infantil. *Anuario de Psicología Jurídica*, 24, 57-63. <https://doi.org/10.1016/j.apj.2014.08.001>
- Vázquez, C. (2015). *De la prueba científica a la prueba pericial*. Marcial Pons.
- Vázquez, C. y Fernández López, M. F. (2022). La valoración de la prueba I: La valoración individual de la prueba. En J. Ferrer Beltrán (Coord.), *Manual de razonamiento probatorio* (pp. 289-351). Suprema Corte de Justicia de México.
- Wells, G. L. (1984). How Adequate is Human Intuition for Judging Eyewitness Memory? En G. L. Wells y E. Loftus (Comp.), *Eyewitness Testimony. Psychological Perspectives* (pp. 256-272). Cambridge University Press.
- Wells, G. L. y Lindsay, R. C. L. (1983). How Do People Infer the Accuracy of Eyewitness Memory? Studies of Performance and Metamemory Analysis. En S. M. A. Lloyd-Bostock y B. R. Clifford (Comp.), *Evaluating Witness Evidence* (pp. 41-55). John Wiley & Sons.
- Wise, R. A., Pawlenko, N. B., Safer, M. A. y Meyer, D. (2009). What US Prosecutors and Defense Attorneys Know and Believe about Eyewitness Testimony. *Applied Cognitive Psychology*, 23(9), 1266-1281. <https://doi.org/10.1002/acp.1530>
- Wise, R. A. y Safer, M. A. (2004). What US Judges Know and Believe about Eyewitness Testimony. *Applied Cognitive Psychology*, 18(4), 427-443. <https://doi.org/10.1002/acp.993>
- Yarmey, A. D. y Jones, H. P. T. (1983). Is the Psychology of Eyewitness Identification a Matter of Common Sense? En S. M. A. Lloyd-Bostock y B. R. Clifford (Comp.), *Evaluating Witness Evidence* (pp. 13-40). Wiley.

Bibliografía complementaria

- Echeburúa, E., Muñoz, J. M. y Loinaz, I. (2011). La evaluación psicológica forense frente a la evaluación clínica: propuestas y retos de futuro. *International Journal of Clinical and Health Psychology*, 11(1), 141-159. https://www.aepc.es/ijchp/articulos_pdf/ijchp-375.pdf

- Herrero, C. (2021). Los conocimientos de la psicología más allá de la prueba pericial. *Quaestio facti. Revista Internacional Sobre Razonamiento Probatorio*, 2021(2), 363-408. <https://raco.cat/index.php/quaestio-facti/article/view/399970>.
- Melinder, A., Goodman, G. S., Eilertsen, D. E. y Magnussen, S. (2004). Beliefs about Child Witnesses: A Survey of Professionals. *Psychology, Crime & Law*, 10(4), 347-365. <https://doi.org/10.1080/10683160310001618717>

Legislación y jurisprudencia

- Anteproyecto de Ley de Enjuiciamiento Criminal (2020) <https://www.mjusticia.gob.es/es/AreaTematicalActividadLegislativa/Documents/210126%20ANTEPROYECTO%20LECRIM%202020%20INFORMACIÓN%20PUBLICA%20%281%29.pdf>
- España. Ley 4/2015, de 27 de abril, del Estatuto de la víctima del delito. (BOE, núm. 101, 28/04/2015). <https://www.boe.es/eli/es/l/2015/04/27/4/con>

ANEXO

Cuestionario sobre creencias sobre la exactitud de los testimonios. Adaptado de Kassin *et al.* (2001)

1. Niveles muy altos de estrés perjudican la precisión de las declaraciones de los testigos.	V
2. La presencia de un arma perjudica la capacidad del testigo para identificar con precisión al autor de los hechos.	V
3. El uso de una rueda de reconocimiento con una sola persona en vez de varias incrementa el riesgo de identificaciones erróneas/incorrectas	V
4. Cuanto más se parezcan los miembros de la rueda de reconocimiento, mayor será la probabilidad de que la identificación del sospechoso sea precisa.	V
5. Las instrucciones de la Policía pueden afectar la voluntad de un testigo al hacer una identificación.	V
6. Cuanto menos tiempo tenga un testigo para observar un evento, menos lo recordará en detalle.	V
7. La tasa de pérdida de la memoria de un evento es mucho mayor justo después del hecho y luego se nivela con el tiempo.	V
8. La confianza de un testigo no es un buen predictor de la precisión de su capacidad de identificación.	V
9. El testimonio de los testigos acerca del evento a menudo refleja no solo lo que realmente vieron, sino también la información que recibieron después.	V
10. Los juicios realizados sobre un color bajo una luz monocromática (ej: luz de farola anaranjada/ámbar) son muy poco fiables.	F
11. El testimonio de un testigo acerca de un evento puede verse afectado por la forma en que se formulan las preguntas al testigo.	V
12. El testigo a veces identifica como culpable a alguien que ha visto en otras situaciones	V

13. Los agentes de Policía y otros observadores entrenados no son más precisos como testigos que un ciudadano medio.	V
14. La hipnosis aumenta la precisión de la memoria de un testigo.	F
15. La hipnosis incrementa la sugestión a las preguntas capciosas y engañosas.	V
16. La percepción y la memoria del testigo de un evento pueden verse afectadas por sus actitudes y expectativas.	V
17. Los testigos tienen mayor dificultad en recordar eventos violentos que no violentos.	V
18. Los testigos son más precisos cuando identifican a personas de su propia raza que a personas de otras.	V
19. La confianza de un testigo puede verse influida por factores que no están relacionados con la precisión de la identificación.	V
20. La intoxicación alcohólica afecta a la capacidad posterior del testigo para recordar personas y eventos.	V
21. La exposición a las reseñas policiales (fotografías) de un sospechoso aumenta la probabilidad de que el testigo reconozca posteriormente a ese sospechoso en la rueda de reconocimiento.	V
22. Las experiencias traumáticas pueden ser reprimidas durante muchos años y después recuperarse.	F
23. Los recuerdos de las personas con relación a su infancia son a menudo falsos o están distorsionados de alguna manera.	V
24. Es posible diferenciar de forma fiable entre recuerdos verdaderos y falsos.	F
25. Los niños pequeños son menos precisos que los testigos adultos.	V
26. Los niños pequeños son más vulnerables que los adultos a las recomendaciones del entrevistador, las presiones de los compañeros y otras influencias sociales.	V
27. Cuanto más se parezcan los miembros de la rueda de reconocimiento a la descripción del culpable aportada por el testigo, más probable será que la identificación del sospechoso sea precisa.	V
28. Es más probable que los testigos identifiquen erróneamente a alguien al hacer un juicio relativo cuando se les presente una rueda simultánea (en lugar de secuencial). Una rueda simultánea es aquella en la que todos los miembros de la rueda de reconocimiento se presentan al mismo tiempo, y se le pregunta al testigo si el culpable o sospechoso es alguno de ellos. En una rueda de reconocimiento secuencial, se presenta a quienes participan en la rueda de uno en uno y se pregunta al testigo si esa persona es culpable/sospechosa.	V
29. Los testigos ancianos son menos precisos que los adultos más jóvenes.	V
30. Cuanto más rápido identifica un testigo en una rueda de reconocimiento, más preciso será.	V

Qf



Marcial
Pons



Càtedra
de Cultura
Jurídica



Universitat
de Girona



Diputació de Girona

www.quaestiofacti.com