Quaestio facti

Revista Internacional sobre Razonamiento Probatorio
International Journal on Evidential Legal Reasoning

2026 | N. 11

ISSN: 2604-6202

ENSAYOS

Miguel Katz Zagury Fragelli, *Prova pericial e vieses cognitivos: o problema dos quesitos e dos quesitos complementares no processo penal brasileiro*

Renzo Cavani, *«Carga» de la prueba: del réquiem a la obertura*

Sebastián Bravo Ibarra, Eduardo Estrada Aravena, *Más allá de la interfaz: deslindes y singularidades del* truthmaker digital

Piero Mattei-Gentili, *Proving Customary Norms:Ontology and Evidence in International Law*

Nathalie Mousist, *¿Ven lo mismo expertos y novatos? Sobre la estructura de la experiencia perceptualy la fuerza de las inferencias*

CONJETURAS Y REFUTACIONES

Silvia Bozza, Franco Taroni, Colin Aitken, *In Defence of Subjectivity: Extending the Argument for a Transparent Trial. (A Support on "Bayesian modelling of criminal cases as a whole. A philosophical reflection on Dutch case law")*

Marcello Di Bello, *Events and Identity in Probabilistic Models of legal Evidence*

Mario Günther, *Bayesian Decision Theory Can Guide Legal Factfinding*

Leya Lisa Hampson, *Combatting Bayesian Criticism: A Bayesian-inspired Critical Checklist for Judicial Reasoning*

Bojan Spaić, *What Goes Around Comes Around: Why the Critique of Expert-Led Integral Bayesian Modelling Also Undermines Judge-Led Bayesian Modelling*

Càtedra de Cultura Jurídica | Universitat de Girona

Marcial Pons

Quaestio *f*acti

Revista Internacional sobre Razonamiento Probatorio
International Journal on Evidential Legal Reasoning

2026 | N. 11

www.quaestiofacti.com

Càtedra de Cultura Jurídica | Universitat de Girona

Marcial Pons

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

Carmen Vázquez
Universitat de Girona, España

COMITÉ DE REDACCIÓN

Daniela Accatino
Universidad Austral de Chile, Chile

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

Jordi Ferrer
Universitat de Girona, España

Daniel González Lagier
Universidad de Alicante, España

Rachel Herdy
Universidade Federal do Rio de Janeiro

Florencia Rimoldi
Universidad de Buenos Aires

SECRETARÍA DE REDACCIÓN Y EDICIÓN

Anna de Castro Cros
Universitat de Girona, España

COMITÉ EDITORIAL

Silvia Bozza
Ca' Foscari University of Venice, Italia

Flavia Carbonell
Universidad de Chile, Chile

Christian Dahlman
Lund University, Suecia

Mauricio Duce
Universidad Católica de Chile, Chile

Mercedes Fernández
Universidad de Alicante, España

Raymundo Gama
Instituto Tecnológico Autónomo de México

Marina Gascón
Universidad de Castilla-La Mancha, España

Clarissa Diniz Guedes
Universidade Federal de Juiz de Fora, Brasil

Ho Hock Lai
National University of Singapore, Singapur

Anne Ruth Mackor
University of Zurich, Suiza

Antonio Manzanero
Universidad Complutense de Madrid, España

Paul Roberts
University of Nottingham, Reino Unido

Marta Saad
Universidade de São Paulo, Brasil

Sarah Summers
University of Zurich, Suiza

Giovanni Tuzet
Università Bocconi, Italia

Jonatan Valenzuela
Universidad de Chile, Chile

CONSEJO ASESOR

Ronald Allen
Northwestern University, EEUU

Amalia Amaya
Universidad Nacional Autónoma de México, México

Lorena Bachmaier
Universidad Complutense de Madrid, España

Zhang Baosheng
University of Political Science and Law (CUPL), China

Lorenzo Bujosa
Universidad de Salamanca, España

Rodrigo Coloma
Universidad Alberto Hurtado, Chile

Elena d'Alessandro
Università di Torino, Italia

Jacopo Della Torre
Università di Genova, Italia

Margarita Diges
Universidad Autónoma de Madrid, España

Gary Edmond
University of New South Wales, Australia

Olga Fuentes Soriano
Universidad Miguel Hernández

Giulia Lasagni
Università di Bologna, Italia

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

Vitor Lia de Paula Ramos
Pontifícia Universidade Católica do Rio Grande do Sul, Brasil

Paula Pessoa
Universidad de Brasília, Brasil

Catherine Piché
Université de Montréal, Canadá

Joan Picó i Junoy
Universitat Pompeu Fabra, España

Giovanni Priori
Pontificia Universidad Católica del Perú, Perú

Maria Thereza Rocha de Assis Moura
Universidade de São Paulo, Brasil

Ana Rodríguez Álvarez
Universidad de Santiago de Compostela, España

Paulo de Sousa Mendes
Universidade de Lisboa, Portugal

Giulio Ubertis
Università Cattolica del Sacro Cuore, Italia

Alan Uzelac
University of Zagreb, Croacia

Susan Haack †
Michele Taruffo †

ÍNDICE

ENSAYOS

Miguel Katz Zagury Fragelli, *Prova pericial e vieses cognitivos: o problema dos quesitos e dos quesitos complementares no processo penal brasileiro* 11

Renzo Cavani, *«Carga» de la prueba: del réquiem a la obertura* 45

Sebastián Bravo Ibarra, Eduardo Estrada Aravena, *Más allá de la interfaz: deslindes y singularidades del* truthmaker digital..... 63

Piero Mattei-Gentili, *Proving Customary Norms:Ontology and Evidence in International Law* 89

Nathalie Mousist, *¿Ven lo mismo expertos y novatos? Sobre la estructura de la experiencia perceptualy la fuerza de las inferencias* 123

CONJETURAS Y REFUTACIONES

Silvia Bozza, Franco Taroni, Colin Aitken, *In Defence of Subjectivity: Extending the Argument for a Transparent Trial. (A Support on "Bayesian modelling of criminal cases as a whole. A philosophical reflection on Dutch case law")* 143

Marcello Di Bello, *Events and Identity in Probabilistic Models of legal Evidence*..... 157

Mario Günther, *Bayesian Decision Theory Can Guide Legal Factfinding*..... 181

Leya Lisa Hampson, *Combatting Bayesian Criticism: A Bayesian-inspired Critical Checklist for Judicial Reasoning*..... 199

Bojan Spaić, *What Goes Around Comes Around: Why the Critique of Expert-Led Integral Bayesian Modelling Also Undermines Judge-Led Bayesian Modelling* 225

ENSAYOS

Quaestio facti. Revista Internacional sobre Razonamiento Probatorio
Quaestio facti. International Journal on Evidential Legal Reasoning
Sección: Ensayos
2026 I 11 pp. 11-44
Madrid, 2026
DOI: 10.33115/udg_bib/qf.i11.23228
Marcial Pons Ediciones Jurídicas y Sociales
© Miguel Katz Zagury Fragelli
ISSN: 2604-6202
Recibido 1/11/25 | Aceptado 25/5/26 | Publicado *online*: 05/06/2026
Editado bajo licencia Reconocimiento 4.0 Internacional de Creative Commons

PROVA PERICIAL E VIESES COGNITIVOS: O PROBLEMA DOS QUESITOS E DOS QUESITOS COMPLEMENTARES NO PROCESSO PENAL BRASILEIRO

Miguel Katz Zagury Fragelli*

RESUMO: O artigo aborda o problema dos vieses cognitivos na produção da prova pericial, focando no processo penal brasileiro. A partir da análise descritiva de estudos empíricos realizados, conceitua esses vieses, no que consistem os vieses de confirmação e de contexto, como podem afetar essa prova e o que fazer para limitá-los, analisando se o princípio do contraditório pode servir para tanto. Assim, examina se esses vieses podem afetar a fase de formulação de quesitos — iniciais ou complementares — aos peritos, uma etapa inicial na formação dessa prova, mas que pode colocar em risco seus resultados. Ao final, a partir do tema dos quesitos, pretende demonstrar o risco de tratar a prova pericial apenas por regras processuais que não prestam atenção ao problema dos vieses cognitivos, o que pode significar que institutos desenvolvidos para melhorar a qualidade das perícias somente aumentem o risco de ocorrência desses vieses.

* Mestre em Raciocínio Probatório pela Universitat de Girona e pela Università degli Studi di Génova. Especialista em Direito Penal Econômico pela Fundação Getúlio Vargas (FGV-SP). Graduado em Direito pela Pontifícia Universidade Católica de São Paulo (PUC-SP). Advogado.

Esse artigo é uma versão atualizada do *Trabajo Final del Máster* que apresentei no programa do Mestrado em Raciocínio Probatório da Universitat de Girona e da Università degli Studi di Génova. Nesse trabalho, tive o privilégio e o prazer de ser orientado pela professora Carmen Vázquez, a quem agradeço por todos os comentários e sugestões feitos, sem os quais o trabalho certamente teria um resultado muito diferente. Agradeço também os professores Vitor de Paula Ramos e Edgar Aguillera, membros da banca que avaliou o trabalho, pelos comentários e perguntas que me instigaram reflexões incorporadas nesta versão. Por fim, não poderia deixar de agradecer os amigos que fiz durante o Mestrado, com os quais discuti as primeiras ideias deste artigo: Matheus Badaró, Luis Felippe Fonseca Católico, Renato Gandini, Daniel Mobley Grillo e Moacyr Coimbra.

PALAVRAS-CHAVE: processo penal; prova pericial; vieses cognitivos; princípio do contraditório; quesitos.

EXPERT WITNESS AND COGNITIVE BIASES: THE PROBLEM OF THE QUESTIONS AND THE COMPLEMENTARY QUESTIONS IN THE BRAZILIAN CRIMINAL PROCEDURE

ABSTRACT: The article addresses the problem of cognitive biases in the production of expert evidence, focusing on the Brazilian criminal procedure. Based on a descriptive analysis of empirical studies, it conceptualizes these biases, what confirmation and contextual biases consist of, how they can affect this evidence and what could be done to limit them, analyzing if the adversarial principle can serve this purpose. It examines whether these biases can affect the phase of formulating questions — initial or complementary — to experts, an initial stage in the formation of expert evidence, but that can compromise its results. Finally, based on the theme of those questions, it aims to demonstrate the risk of treating expert evidence through procedural rules that do not pay attention to the problem of cognitive bias, which can mean that institutes developed to improve the quality of expert evidence only increases the risk of these biases occurring.

KEYWORDS: criminal procedure; expert evidence; cognitive biases; adversarial principle; questions.

SUMÁRIO: 1. INTRODUÇÃO.— 2. PROVA PERICIAL E VIESES COGNITIVOS: RISCOS E CONTROLES: 2.1. Vieses cognitivos na prova pericial; 2.2. Contraditório, parcialidade cognitiva e controle efetivo da prova pericial.— 3. PRODUÇÃO DA PROVA PERICIAL NO PROCESSO PENAL BRASILEIRO: 3.1. Contexto da produção da prova pericial; 3.2. Rito de produção da prova pericial; 3.3. Caso de estudo: quesitos e quesitos complementares.— 4. QUESITOS, QUESITOS COMPLEMENTARES E PROBLEMAS RELATIVOS AOS VIESES COGNITIVOS NA PRODUÇÃO DA PROVA PERICIAL NO PROCESSO PENAL BRASILEIRO: 4.1. Quesitos no inquérito policial: possível fonte de vieses de confirmação e de contexto; 4.2. Quesitos complementares: possível fonte de viés de confirmação.— 5. CONCLUSÃO.— Bibliografia.

1. INTRODUÇÃO

Até a segunda metade do século XX, a maioria das investigações criminais baseava-se apenas em provas testemunhais ou relacionadas aos instrumentos ou proveitos do delito, sem utilizar provas periciais (Reid, 2018, p. 106). Atualmente, essa tendência foi alterada.

Recentes inovações tecnológicas atribuíram grande importância à prova pericial, pois permite a comprovação de fatos que antes dificilmente eram demonstrados por outros meios (Gascón Abellán, 2013, p. 181-182).

A prova pericial tornou-se dominante, sendo a que mais influencia o resultado dos casos (Duce, 2013, p. 16-17). É uma circunstância que decorre do desenvolvimento da sociedade, não do desenho dos sistemas jurídicos (Duce, 2022, p. 148-149).

No processo penal, objeto deste trabalho, isso é fortalecido pela tipificação, em diferentes ordenamentos, de crimes que tutelam bens jurídicos coletivos, como o meio ambiente e o sistema financeiro (Tonini, 2004, p. 196; Malan, 2015, p. 293). Assim, essa prova é necessária para explicar como o fato ocorreu e o que o causou (Tonini, 2004, p. 196) e para explicar ao juiz e às partes questões que não são de seu domínio ou conhecimento (Malan, 2015, p. 293).

Trata-se da finalidade da prova pericial: introduzir conhecimento especializado ao processo. Como isso ocorre por meio de um ato de comunicação em que um terceiro fornece um conhecimento a quem ele informa, a prova pericial é, em termos epistemológicos, um testemunho (Vázquez, 2021, p. 82-88).

Assim, este trabalho adota três posições epistemológicas que orientam seu desenvolvimento. A primeira é a de que, tal qual uma testemunha comum, o perito oferece uma interpretação sobre os fatos, havendo uma diferença de grau, mas não de categoria, entre ambos (Vázquez, 2021, p. 88-89). Enquanto qualquer pessoa pode ser uma testemunha, um perito precisa ter conhecimento teórico e prático: deve ter formação técnica em determinada área e saber como utilizá-la quando solicitado (Vázquez, 2021, p. 71-72).

A segunda é a de que, no processo, a principal finalidade da prova é a busca pela verdade, sendo um meio para obter conhecimento sobre os fatos (Ferrer Beltrán, 2017, p. 58).

A terceira é a de que, para esse processo ser racionalmente controlável, a verdade deve ser compreendida pela noção de correspondência: um enunciado sobre os fatos será verdadeiro quando representar o que realmente ocorreu (Badaró, 2019, p. 86-88).

Logo, o objeto de um processo não consiste nos fatos em si, mas sim no enunciado fático, na hipótese formulada sobre esses fatos (Badaró, 2019, p. 70-71). Afinal, todo processo de conhecimento de um fato é influenciado pela interpretação de quem o descreve (Prado, 2024, p. 390).

Mesmo que o objetivo primário da prova seja a obtenção da verdade, trata-se de um meio que possui limitações e que não garante por si só que a verdade será alcançada. Assim, o enunciado fático será dado por provado não quando for verdadeiro, mas sim quando existirem, no processo, elementos que suficientemente corroborem a hipótese (Ferrer Beltrán, 2017, p. 34-39).

Portanto, a prova pericial tem a função de fornecer elementos para que os operadores do direito possam determinar se certo enunciado fático corresponde à realidade, tendo a particularidade de tratar sobre temas que não são de domínio do juiz e das partes. É a mesma lógica da prova testemunhal, mas com uma diferença: nas perícias, a informação tem que ser transmitida por um especialista em certa área do

conhecimento, ao passo que, em testemunhos comuns, isso é feito por um leigo, via exercício de memória.

Apesar de essas provas pertencerem à mesma categoria epistemológica, os operadores do direito costumam considerá-las de formas diferentes. Em relação às perícias, costuma-se entender que a análise de evidências é um processo objetivo, independente e infalível (Dror, Charlton, Péron, 2006, p. 74; Edmond *et al*, 2015, p. 02).

Apresentada com caráter científico, a prova pericial não recebe um tratamento mais crítico e cuidadoso pelos sistemas processuais, que costumam concebê-la como uma verdade absoluta. Ao tratar das perícias que fazem uso de ciências forenses, Marina Gascón Abellán discorre sobre a sobrevalorização da perícia, frequentemente entendida como um meio que com certeza produzirá uma decisão correta (Gascón Abellán, 2013, p. 182).

Tal sobrevalorização possui tanto uma dimensão epistêmica quanto semântica. A primeira refere-se ao tratamento das perícias como inquestionáveis, enquanto a última relaciona-se com a falha de interpretação que muitas vezes acompanha esse meio de prova —geralmente, entende— se que uma prova pericial vincula um vestígio a uma única fonte, sendo que deveria apenas apontar um grau de probabilidade (Gascón Abellán, 2013, p. 182-187).

Essa concepção mais difundida, chamada de paradigma da individualização, faz com que o perito seja quem determine o que se crê sobre um vestígio —se é proveniente ou não de uma pessoa, taxativamente—, o que é uma tarefa do juiz (Gascón Abellán, 2013, p. 189). Assim, exige-se que o perito faça uma valoração das hipóteses fáticas em seus exames.

O ideal seria que essa prova fosse tratada pelo paradigma da verossimilhança. O perito deve indicar o grau de probabilidade dos dados que examina à luz de todas as hipóteses formuladas, não o grau de probabilidade das hipóteses a partir desses dados (Gascón Abellán, 2013, p.194).

Essa situação das provas periciais decorrentes de ciências forenses exemplifica o tratamento pouco cuidadoso que geralmente se confere às perícias. Não é algo irrelevante, embora muito ignorado. Em razão disso, podem ser produzidos exames periciais de baixa qualidade e interpretações equivocadas do significado dessas provas, o que tem grande potencial de causar decisões judiciais equivocadas.

Diferentemente do que ocorre na maioria dos ordenamentos jurídicos, é necessário que sejam mitigados os riscos inerentes a esse meio de prova. É nesse contexto que este trabalho visa analisar se o processo penal brasileiro está capacitado para lidar com um desses riscos: a contaminação de exames periciais por vieses cognitivos, compreendidos neste trabalho como erros sistemáticos incidentes no raciocínio ao processar e interpretar informações.

Será feita uma análise descritiva para demonstrar, por estudos empíricos já realizados, o que são esses vieses, no que consistem os vieses cognitivos de confirmação e de contexto e como podem incidir na produção de perícias.

Com base nessa análise, o trabalho tem, como objetivo mais geral, examinar se esses vieses podem incidir na produção de perícias no processo penal brasileiro e qual é o papel que o contraditório pode ter para mitigar riscos decorrentes da contaminação cognitiva.

Para deixar essa análise mais concreta, o trabalho utilizará como caso de estudo a etapa de formulação de quesitos — iniciais ou complementares—, que, por ter um papel muito inicial na conformação das perícias, pode comprometer seus resultados.

2. PROVA PERICIAL E VIESES COGNITIVOS: RISCOS E CONTROLES.

Embora tida como uma fonte de certeza absoluta, sendo cada vez mais utilizada, a prova pericial é um dos principais fatores que causam condenações de inocentes (Duce, 2013, p. 47-49; Duce, 2022, p. 145, 151-152). No processo penal brasileiro, essa afirmação pode soar um tanto abstrata, dada a ausência de dados empíricos a respaldá-la.

Contudo, a experiência comparada é suficientemente sólida[1] para apontar que os riscos da prova pericial são inerentes a todos os sistemas de justiça penal (Duce, 2013, p. 51). Portanto, mesmo diante da falta de estudos empíricos específicos, trata-se de tema que deve causar preocupação no Brasil.

Um país que pode ser citado como exemplo é os Estados Unidos da América. Em consulta ao registro nacional de condenações equivocadas revertidas, verifica-se que, das 3.586 exonerações contabilizadas desde 1989, 1.029 (28,69%) estavam relacionadas ao mau uso de ciências forenses[2]. Nesse sentido, após analisar 250 condenações equivocadas que resultaram em exonerações, o pesquisador Brandon Garrett verificou que em 128 (51,20%) houve algum problema relacionado à prova pericial (Garrett, 2011, p. 114)[3].

Em estudo realizado pelo National Research Council, foi constatado que, exceção feita aos exames de DNA, as provas periciais utilizam ciências forenses que não possuem a objetividade e a validade científica que imaginamos ter (National Research Council, 2009, p. 07). Diante disso, não surpreende que ciências forenses inválidas e sem fiabilidade empírica estejam muito associadas a condenações de inocentes (Garrett, 2011, p. 89-91).

---

[1] Para maiores dados, referentes a outros países, ver Duce, 2013, p. 46-51.

[2] Disponível em: https://www.law.umich.edu/special/exoneration/Pages/Exonerations-in-the-United-States- Map.aspx. Acesso em 06/04/2025, 14h35.

[3] Sobre as condenações equivocadas causadas por problemas de provas periciais, tem-se que isso não é um fator excludente. Não raras vezes também se verifica a presença de outros fatores, como confissões falsas.

Com a ausência de objetividade, há grande espaço para interpretação de quem realiza o exame pericial. Quanto mais espaço há para subjetividade, maiores são as chances de o perito ser contaminado por um viés cognitivo que influencie a sua análise (Garrett, 2021, p. 116).

Logo, os vieses cognitivos afetam a imparcialidade dos peritos, causando o que Carmen Vázquez chama de parcialidade cognitiva: a contaminação de perícias pelas predisposições cognitivas dos peritos e/ou informações com as quais têm contato (Vázquez, 2021, p. 132)[4].

Cognitivamente comprometido, o perito pode apontar resultados que não correspondam à realidade. Como poder fazer que o que esteja provado não seja verdadeiro, os vieses cognitivos são um dos fatores que levam perícias a causarem erros judiciais (Duce, 2022, p. 163-165).

Assim, o primeiro objetivo do trabalho é analisar a contaminação de perícias no processo penal por vieses cognitivos e a necessidade de termos sistemas processuais estruturados para impedir que ocorram (Páez, 2021, p. 188)[5].

2.2 Vieses cognitivos na prova pericial

Principalmente em sistemas processuais penais acusatórios, costuma-se haver preocupação apenas com peritos parciais em favor da parte que solicitou suas análises (Edmond *et al*, 2015. p. 22). Afinal, peritos tendem a ficar enviesados pela parte que os demanda, seja a defesa ou a acusação (Garrett, 2021, p. 113-114). É uma parcialidade de origem (Vázquez, 2021, p. 132)[6].

Na maioria das vezes, essa preocupação restringe-se às provas periciais produzidas pela defesa, sem muita atenção sobre a qualidade das perícias produzidas pelo Estado e que tenham carga incriminatória (Edmond *et al*, 2015, p. 23). Nesse contexto, devemos prestar atenção à parcialidade cognitiva, que afeta a qualidade da prova e, diferentemente da parcialidade de origem, é de difícil identificação (Vázquez, 2021, p. 135).

Essa parcialidade é causada por vieses cognitivos. São erros sistemáticos que ocorrem no raciocínio de uma pessoa ao processar e interpretar informações, o que afeta as conclusões tomadas (Vázquez, 2022a, p. 79). Em matéria pericial, os vieses cogni-

---

[4] Neste artigo, o problema da parcialidade cognitiva será analisado apenas no que se refere aos vieses de confirmação e de contexto.

[5] A limitação — idealmente, eliminação — dos vieses cognitivos não se resume ao objetivo de ter provas com maior qualidade. É necessária para termos decisões judiciais racionalmente legitimadas. Se os vieses são elementos que deslegitimam decisões, é necessário existirem regras que limitem seus efeitos. Sobre o tema, ver Páez, 2021.

[6] Costuma haver uma confusão sobre todo perito de parte ser parcial em favor de quem o contratou. Isso não é verdade: mesmo que a decisão sobre a juntada dessa perícia pertença à parte, que só juntará o que lhe favorecer, isso não significa que o perito tenha conduzido seus exames para beneficiá-la (Vázquez, 2021, p. 133).

tivos podem fazer com que o perito ignore certas evidências, supervalorize outras e mude seus critérios de decisão, chegando a uma conclusão equivocada (Dror, Charlton, Péron, 2006, p. 77).

Tais vieses impactam como os dados são coletados, observados e compreendidos, como os resultados são interpretados e como as conclusões são tomadas (Dror, 2020, p. 7998), e, portanto, podem fazer com que a perícia indique algo que não corresponda à realidade.

A maioria das provas periciais é fundamentada em ciências forenses que dependem da interpretação de uma pessoa para individualizar a origem de um vestígio a uma pessoa. Só que essa interpretação pode ser contaminada por informações e condições existentes na atividade do perito (Edmond *et al*, 2015, p. 10-11). Logo, os vieses cognitivos podem fazer com que se chegue a uma conclusão equivocada quanto à individualização de um vestígio, o que pode causar erros judiciais.

O problema disso é que esses vieses incidem de uma forma que os peritos não percebem o comprometimento do seu raciocínio, o que os torna ainda mais perigosos (Giannelli, 2010, p. 257; Dror, 2020, p. 7999).

Nas perícias, os principais vieses cognitivos são os de contexto e de confirmação[7], que podem incidir em todo tipo de prova pericial (Vázquez, 2023, p. 31). Portanto, os sistemas processuais deveriam estar estruturados para tentar evitá-los, sendo necessário saber como podem acontecer[8].

O viés contextual ocorre quando uma informação irrelevante para a perícia altera a expectativa do examinador, alterando o que percebe sobre o que examina e, assim, a sua conclusão (Giannelli, 2010, p. 253; Vázquez, 2022a, p. 82). São informações que, embora relevantes para o caso, são irrelevantes para o raciocínio pericial (Vázquez, 2023, p. 31).

Esse tipo de viés pode ocorrer de diferentes formas, como quando o mesmo perito é responsável por coletar os vestígios e posteriormente analisá-los, pois poderá ser contaminado por elementos presentes na cena do crime, mas que não importam para o raciocínio pericial (Vázquez, 2022a, p. 84).

Informações sobre os antecedentes criminais de uma pessoa ou sobre o suspeito ter confessado também podem causar tal contaminação (National Commission on Forensic Science, 2015, p. 03). São informações irrelevantes, classificadas como as que em nada contribuem para determinar a probabilidade de uma evidência ser proveniente ou não de certa pessoa (National Commission on Forensic Science, 2015, p. 03).

---

[7] Um levantamento de diversos estudos que comprovam empiricamente a ocorrência desses vieses é encontrado em Cooper, Meterko, 2019, que foca em provas periciais que fazem uso de ciências forenses.

[8] Por ser uma análise do processo penal brasileiro, estre trabalho foca em soluções institucionais para lidar com os vieses cognitivos, à luz do quanto defendido por Páez, 2021. Contudo, para evitar ao máximo a ocorrência desses vieses, também é necessário pensar em soluções individuais, como defendido por Amaya, 2022.

Por sua vez, o viés confirmatório acontece quando crenças e expectativas prévias do examinador influenciam as atividades de colheita, percepção e interpretação das evidências (Kassin, Dror, Kukucka, 2013, p. 45), fazendo com que oriente sua análise para confirmá-las (Edmond *et al*, 2015, p. 06). Logo, o perito testará uma hipótese visando exclusivamente a sua corroboração por meio de um raciocínio circular, sem buscar informações que poderiam rejeitá-la (Giannelli, 2010, p. 254).

As crenças de um indivíduo costumam causar esse tipo de viés por meio de três etapas: primeiro, fazem que o examinador crie uma impressão sobre o que está analisando; depois, o examinador adota uma postura coerente com essa impressão; ao final, o que está sendo analisado é ajustado para se adequar à impressão (Kassin, Dror, Kukucka, 2013, p. 44-45).

Essas crenças e expectativas podem ser geradas por diferentes fatores. Em artigo no qual reúnem diversos estudos empíricos e conceituam o que seria o viés de confirmação forense, Saul Kassin, Itiel Dror e Jeff Kukucka demonstram que isso pode ser causado pelo papel que os peritos acreditam ter, pela situação em que estão inseridos e até por informações que também podem causar um viés de contexto, como confissões do suspeito (Kassin, Dror, Kukucka, 2013).

O que os diferencia é que, enquanto a contaminação que ocorre no viés contextual é causada por uma informação irrelevante que altera a interpretação do perito, os efeitos do viés confirmatório são gerados por expectativas do examinador sobre o que irá analisar. A despeito disso, ambos aumentam o risco de que os peritos ignorem hipóteses alternativas e interpretem as evidências a fim de corroborar certa hipótese.

Uma das formas mais clássicas de causar isso é indicar, desde o começo da investigação, um suspeito principal. Mas sequer é necessário que exista um suspeito perseguido: basta que exista uma teoria, um padrão ou qualquer informação que faça com que se escolha quais elementos serão analisados, ignorando outros. Assim, essas informações podem não só causar o viés de contexto, mas também o de confirmação (Dror, 2020, p. 8001).

Diferentemente do que o senso comum sugere, ser especialista em certa matéria não torna uma pessoa imune aos vieses cognitivos, mas sim mais suscetível (Dror, 2020, p. 7999). O próprio processo de especialização faz com que valorize alguns dados e confie mais em determinadas expectativas em razão de raciocínios prévios, o que a torna vulnerável a processar incorretamente alguma informação (Dror, 2020, p. 7999; Tversky, Kahneman, 1974, p. 1131).

Até porque, considerando que os vieses cognitivos podem ser causados por princípios heurísticos utilizados para simplificar raciocínios complexos (Tversky, Kahneman, 1974, p. 1124), o processo de especialização pode sedimentar essas heurísticas com maior facilidade nos especialistas.

Erros cometidos por causa desses vieses podem ser feitos por pessoas bem-intencionadas (Dror, Cole, 2010, p. 162) e não têm qualquer relação com a competência

delas (Dror, 2020, p. 7999). Até os mais experientes peritos podem sofrer vieses cognitivos e causar sérios impactos (Edmond *et al*, 2015, p. 23-24).

Portanto, a parcialidade cognitiva não tem apenas o problema de ser de identificação mais difícil que a parcialidade de origem, objeto de tanta preocupação pelos sistemas processuais. Erros causados por vieses cognitivos são mais perigosos porque, por serem cometidos por especialistas que acreditam no que afirmam, são mais persuasivos (Dror, Cole, 2010, p. 162).

Peritos cognitivamente comprometidos podem causar um efeito de bola de neve e enviesar outras pessoas (Dror, 2020, p. 8003). O resultado de uma prova pericial pode comprometer a produção de outras provas, como testemunhos e perícias posteriores (Edmond *et al*, 2015, p. 02-03). Na revisão de perícias, por exemplo, se o perito revisor souber as conclusões do primeiro exame, pode incorrer em um viés confirmatório (Vázquez, 2022a, p. 81).

Logo, os vieses cognitivos são um grande problema, pois podem ocorrer em qualquer perito, não dependem de uma intenção prévia e possuem relevante grau de persuasão (Vázquez, 2022a, p. 80). Quando se trata do processo penal e das consequências de um erro judicial, isso fica ainda mais grave.

A primeira coisa a ser feita para diminuir os riscos dos vieses é entender os fatores que aumentam as chances de ocorrerem. Não são erros simples, mas comportamentos que tendem a se repetir em circunstâncias parecidas (Páez, 2021, p. 189).

O desenho institucional em que inseridos os órgãos periciais é de grande relevância para maximizar ou minimizar a incidência dos vieses cognitivos (Duce, 2022, p. 165-166). A depender do contexto em que esses órgãos estão alocados, os peritos podem entender que possuem um papel a desempenhar, o que pode influenciar as decisões que tomam (Giannelli, 2010, p. 252).

No processo penal, a prova pericial costuma ser produzida após requerimento do órgão responsável pela investigação ou acusação e a maioria dos laboratórios periciais está inserida na lógica do sistema de combate à criminalidade (Garrett, 2011, p. 92).

Alocá-los nesse sistema aumenta os riscos de que exames sejam contaminados pelo viés confirmatório (Dror, 2020, p. 8002) e pode gerar peritos parciais (Giannelli, 2010, p. 247-250). Peritos que trabalham para a polícia ou para órgãos acusatórios têm maiores chances de ficarem enviesados, pois podem acreditar que fazem parte desse sistema (Garrett, 2021, p. 111-112). Inclusive, não é só a polícia que pode enviesar peritos, o que também pode ser feito pelo Ministério Público (Giannelli, 2010, p. 257-259).

Quando trabalham apenas com essas instituições, peritos podem ter maior tendência a buscar evidências contra o suspeito (Garrett, 2021, p. 109). Podem passar a acreditar que devem fornecer provas incriminatórias, o que pode influenciar seus exames (Garrett, 2011, p. 92). Ficariam mais tendenciosos a terem expectativas desfavoráveis ao suspeito, o que os tornaria mais vulneráveis ao viés confirmatório.

Embora possam acreditar que trabalham para o sistema de combate à criminalidade, a função de um perito no processo penal não é auxiliar a acusação a ganhar o caso, mas sim fornecer conhecimento especializado de forma explicativa (Edmond *et al*, 2016, p. 04). Afinal, sob a concepção racionalista da prova, a perícia é um meio para introdução de informações técnicas, não um instrumento estratégico para vencer o processo.

O ideal seria que esses órgãos ficassem próximos da comunidade acadêmica, podendo realizar exames com o que há de mais avançado em determinada área, mas não é o que geralmente ocorre (Duce, 2022, p. 166). Portanto, a alocação de laboratórios periciais no sistema de combate à criminalidade pode prejudicar a qualidade técnica dos exames e aumentar as chances de que sejam conduzidos de forma cognitivamente parcial.

Uma análise pericial imparcial deve ser feita partindo das evidências ao suspeito. Contudo, se está cognitivamente contaminado, o perito pode interpretar as evidências pelo que já sabe sobre o suspeito (Dror, 2020, p. 8000), o que pode causar conclusões equivocadas e erros judiciais.

Isso é ainda mais preocupante quando considerado que é comum que, durante os exames, peritos fiquem expostos a informações irrelevantes, como perícias sobre outras matérias, o que pode ocasionar um viés de contexto (Edmond *et al*, 2015, p. 11). Não são raras as vezes em que começam a fazer suas análises tendo contato com elementos referenciados como pertencentes ao investigado.

Esses elementos são fontes de contaminação, que podem subverter uma análise pericial: mesmo que devesse ser feita da evidência ao suspeito, passa a ser feita pelo caminho inverso, o que compromete as conclusões sobre a evidência (Dror, 2020, p. 8000-8001).

Para diminuir esse risco, o ideal é que os peritos primeiro analisem as evidências e depois sejam expostos ao material referenciado, o que pode evitar análises feitas do suspeito à evidência (Dror *et al*, 2015, p. 03).

Considerando que informações irrelevantes são um risco, o correto seria restringir o acesso dos peritos a isso (Edmond *et al*, 2015, p. 11; Garrett, 2021, p. 109), bem como a demais fatores de contaminação.

O problema dos elementos referenciados exemplifica como a interação entre órgãos de investigação ou acusação e os laboratórios periciais pode causar vieses cognitivos. Práticas e informações sugestivas podem influenciar os peritos, causando vieses que contaminariam interpretações sobre evidências e conclusões tomadas no exame (Edmond *et al*, 2016, p. 20-21).

A forma pela qual um questionamento é feito a um perito também é capaz de enviesá-lo (Amaral, Bruni, 2023, p. 893-894). Assim, é preciso observar se o questionamento não aumenta os riscos de ocorrência de um viés confirmatório ou contextual que contamine o perito.

Portanto, os vieses cognitivos são um risco para a prova pericial em um processo penal racional e podem desvirtuar a função dessa prova[9]. Se a finalidade da prova no processo é a busca pela verdade e um enunciado fático é verdadeiro quando corresponde ao que aconteceu, a perícia tem a particularidade de ser um meio de obtenção de conhecimento especializado sobre fatos ocorridos em áreas que não são do conhecimento dos operadores jurídicos, que não teriam a capacidade de valorar se o enunciado está corroborado pelas evidências existentes sem o auxílio dessa informação.

Todavia, esses vieses influenciam diretamente como as evidências são percebidas e interpretadas pelo perito e sobre a conclusão emitida após os exames. Portanto, podem fazer que a valoração individual dessa prova seja feita equivocadamente.

Se uma perícia conclui que um vestígio é proveniente de uma pessoa, quando na verdade não é, o juízo sobre o enunciado fático que depende dessa prova muito provavelmente será incorreto – o que, inclusive, demonstra o risco de tratar as provas periciais pelo paradigma da individualização[10]. Sendo assim, pode fazer que a decisão final de um processo não seja correspondente ao que ocorreu, causando erros judiciais.

Os vieses cognitivos podem ser compreendidos como um risco para a qualidade das provas periciais. Uma vez que podem gerar decisões equivocadas, é necessário limitar as chances de acontecerem. Como a ocorrência desses vieses passa muito pelo desenho institucional do sistema, é preciso identificar se as regras processuais podem aumentar ou diminuir essas chances (Páez, 2021, p. 188).

Em um processo judicial, a metodologia de apuração da verdade passa muito pela garantia do contraditório, à qual é atribuído papel fundamental para se chegar a uma decisão final correta. Entende-se que essa garantia tem a capacidade de, pelo estabelecimento de mecanismos de contradição, controlar a qualidade da atividade probatória e propiciar a tomada de uma decisão sobre os fatos que corresponda ao que ocorreu.

Assim, e considerando que é necessário limitar os riscos de ocorrerem vieses cognitivos para termos um processo penal mais racional, é preciso analisar se o contraditório pode contribuir para tanto.

2.2. Contraditório, parcialidade cognitiva e controle efetivo da prova pericial

Processo e contraditório são conceitos intimamente relacionados. É a característica de ser desenvolvido mediante contraditório que distingue o processo judicial de um simples procedimento (Badaró, 2019, p. 36).

---

[9] Por desvirtuar uma prova, os vieses comprometem a sua qualidade. Logo, não basta que uma prova seja lícita: é necessário que tenha sido obtida e produzida por meio de boas práticas forenses (Amaral, Bruni, 2023, p. 895).

[10] Por mais que a adoção do paradigma da verossimilhança não tenha a capacidade de evitar a contaminação de perícias por vieses cognitivos, pode fazer com que os efeitos desses vieses sejam menos graves, uma vez que diminui o valor que é dado às conclusões dos laudos periciais.

Tradicionalmente, compreende-se o contraditório como a garantia que exige que as partes tenham ciência dos atos e termos processuais, bem como a possibilidade de reagir contra esses. Contudo, a partir da evolução da concepção do princípio da igualdade, com a superação do seu paradigma formal para se buscar uma igualdade substancial, passou-se a entender que o contraditório, não pode ser uma mera possibilidade de reação: deve ser uma ferramenta que equilibre a relação de desiguais que há no processo penal (Badaró, 2020, p. 60).

Deve ser um mecanismo que permita a participação efetiva da parte no processo. Nesse sentido, o direito à prova é um aspecto fundamental do contraditório, pois é por meio de provas que as partes podem embasar as suas pretensões e refutar as hipóteses da parte contrária (Badaró, 2019, p. 40-41).

No entanto, como as partes utilizam a prova para satisfazer seus interesses, é possível que não conduzam a atividade probatória de forma neutra, o que prejudica a busca pela verdade (Ferrer Beltrán, 2023, p. 63-64). Se a função principal da prova no processo é ser um meio para se chegar à verdade, uma postura enviesada conflitaria com essa finalidade.

Isso contraria o objetivo do direito de ser um mecanismo que oriente as ações de seus destinatários, o que só pode funcionar se o processo servir para determinar a verdade (Ferrer Beltrán, 2023, p. 51). Para isso não ocorrer, é atribuída grande relevância ao contraditório.

Essa garantia é responsável por impor uma dinâmica dialética ao processo, com a estruturação de um ambiente de ação e reação entre as partes, que poderão apresentar elementos favoráveis e rebater os que lhes forem desfavoráveis (Badaró, 2019, p. 39). Se uma parte não apresentar uma prova prejudicial ao seu interesse, a outra pode fornecê-la, o que enriquecerá o acervo probatório.

O contraditório permite que a valoração das provas não seja feita de forma enviesada, influenciada pela versão de uma única parte e sem outra explicação possível sobre o enunciado fático (Prado, 2024, p. 392). Serve para evitar que apenas existam elementos que corroborem uma única hipótese e que prevaleçam interpretações sobre os fatos que não correspondam à verdade.

Portanto, o contraditório tem uma função cognitiva, sendo uma garantia para a busca da verdade: por meio da contraposição de provas e interpretações sustentadas pelas partes, possibilita um conhecimento maior sobre os fatos, o que qualifica a decisão judicial e diminui a possibilidade de erros (Badaró, 2019, p. 39; Prado, 2024, p. 392).

Em razão dessa finalidade heurística, diz-se que o contraditório é a forma de implementação no processo de ferramentas que facilitam a corroboração de hipóteses pelos elementos de juízo (Ferrer Beltrán, 2023, p. 133).

Dentre tais mecanismos, está a participação das partes na produção da prova, o que tem nítida aplicação em provas pessoais. Em provas que consistem no depoimento de alguém, deve ser permitido o contra-exame do declarante pelas partes, que

poderão fazer perguntas para fortalecer ou questionar a confiabilidade dessa prova para corroborar as suas respectivas hipóteses (Ferrer Beltrán, 2023, p. 133-135).

Contudo, existem provas — como perícias de impressões digitais, de DNA etc. — que, dada a sua natureza, não permitem a intervenção completa das partes na sua produção, o que faz com que, em nome do princípio contraditório, seja permitido que elas proponham provas sobre essa prova para questionar ou confirmar a sua confiabilidade (Ferrer Beltrán, 2023, p. 135-136).

Em matéria de perícias, isso é relevante pois a forma como essa prova é apresentada ao processo pode variar. Em um sistema acusatório, a prova pericial a ser valorada é a declaração do perito em juízo, enquanto em um processo de orientação mais inquisitiva o que é valorado é o laudo pericial escrito (Duce, 2013, p. 38-41). Em uma primeira análise, poderia ser sustentado que isso alteraria o mecanismo que o contraditório incidiria nessa prova: em declarações orais seria via contra-exame e em laudos escritos seria por uma nova prova.

Nesse sentido, autores de países com sistemas processuais penais de orientação mais acusatória tratam o contraditório da prova pericial a partir do contra-exame. Mauricio Duce, por exemplo, afirma que o contra-exame é um instrumento fundamental para a confrontação desse meio de prova. Por meio das perguntas das partes, será possível escrutinar a qualidade da prova e as conclusões veiculadas, bem como identificar falsidades, exageros e parcialidades (Duce, 2014, p. 128).

Mesmo assim, o próprio Duce afirma que, para o contraditório ser efetivo, não se pode reduzir o contra-exame à possibilidade de realizar perguntas: é preciso que existam outras ferramentas de contradição, dentre as quais a possibilidade de produzir uma prova sobre a perícia questionada (Duce, 2014, p. 129).

Assim, neste trabalho, considera-se que o contra-exame também engloba a produção de uma outra prova sobre o laudo pericial. Afinal, ambos servem como mecanismos de contraditório para controlar a prova (Ferrer Beltrán, 2023, p. 134).

O ponto principal é que a confrontação da prova consiste em um elemento central nos sistemas processuais contemporâneos (Duce, 2014, p. 129). Logo, deve ser analisado se o contraditório pode servir para controlar a incidência de vieses cognitivos em perícias.

O melhor momento para se controlar a prova pericial — e evitar erros — é o da produção, pois é quando o perito pode ser questionado a partir de mecanismos fundados no princípio do contraditório (Vázquez, 2022b, p. 117-118).

Essa etapa não é só a de maior relevância para o controle genérico da prova pericial. A produção da prova pericial também é a fase que reúne melhores condições para reconhecer eventual parcialidade cognitiva, justamente por ser a etapa em que é possível questionar o perito sobre suas conclusões (Vázquez, 2021, p. 135).

Dada a difícil identificação da parcialidade cognitiva, as etapas de valoração e justificação da prova podem ter maiores empecilhos para lidar com os vieses cognitivos.

Como envolve a possibilidade de questionar os peritos, a etapa da produção pode permitir que esses vieses sejam identificados e, principalmente, evitados.

O contexto de produção dessa prova também é relevante para o controle via contraditório. Idealmente, se feita na fase processual, permite que as partes acompanhem as atividades periciais, o que lhes possibilita exercer maiores controles durante a execução dos exames e até antes de começarem (Vázquez, 2021, p. 138).

Por outro lado, em fases preliminares do processo penal, há pouco espaço para a defesa questionar as conclusões da prova pericial (Edmond *et al*, 2015, p. 20-21). Na experiência brasileira, não há contraditório formalmente instituído na etapa de investigação, o que dificulta manifestações defensivas sobre a perícia.

No processo penal, as provas periciais geralmente são produzidas no inquérito policial, pois, se houver demora na busca, coleta ou análise de vestígios, será mais difícil de produzir uma prova confiável. Sob o ponto de vista da ciência forense, é justificável que a perícia seja produzida o quanto antes e submetida ao contraditório apenas posteriormente, na fase processual (Amaral, Bruni, 2023, p. 895-896).

Só que isso não é um problema apenas para a defesa, mas também para o objetivo do processo de descobrir a verdade. Como se quer que a decisão final seja tomada a partir de toda a informação disponível, deve ser permitido que desde logo as partes questionem o conteúdo dos exames periciais, mesmo que em fases anteriores à judicial (Vázquez, 2022b, p. 109-110).

Contudo, na maioria dos processos, quando uma prova pericial acusatória é admitida, a defesa não tem como intervir na sua produção desde logo e só pode questioná-la a partir de mecanismos tradicionais do contraditório, como o contra-exame, compreendidos como balizas suficientes para perícias incriminatórias (Edmond *et al*, 2015, p. 18-20).

Trata-se da aplicação do contraditório em sua função de controlar a prova. Todavia, esses mecanismos tradicionais não costumam considerar os efeitos dos vieses cognitivos, sendo impossível considerá-los como ferramentas efetivas para escrutinar a confiabilidade da prova pericial sob o ponto de vista da parcialidade cognitiva (Edmond *et al*, 2015, p. 18-20) [11].

Nessa linha, após analisar 250 condenações equivocadas, Brandon Garrett concluiu que não se pode depender apenas do contra-exame para evitar erros decorrentes de provas periciais (Garrett, 2011, p. 114). Em posições semelhantes, Maurício Duce e Gary Edmond sustentam que não podemos depender apenas do sistema acusatório para impedir esses erros (Duce, 2022, p. 170-171; Edmond, 2020, p. 436-437).

Logo, não surpreende a constatação de que, mesmo quando bem realizado, o contra-exame não elimina os riscos inerentes a uma perícia que faz mal uso da ciência forense (Edmond *et al*, 2016, p. 06).

[11] Ademais, referidos mecanismos de contradição sequer podem ser considerados como boas ferramentas de controle dos demais aspectos da prova pericial (Edmond *et al*, 2015, p. 18-20).

É como se, considerando que a prova pericial pertence ao gênero dos testemunhos, os sistemas processuais confiassem que, assim como serve para controlar a prova testemunhal[12], o contra-exame seria suficiente para balizar perícias. Assim, desconsideram as particularidades dessa espécie.

Trata-se de um problema quando retomado que, segundo a concepção moderna do contraditório, é preciso que existam condições adequadas para as partes reagirem efetivamente. Sem atenção às particularidades da prova pericial, o que não costuma ser feito pelos ordenamentos jurídicos, não há como questionar essa prova e controlar sua confiabilidade (Salaverría, 2021, p. 187-189).

Mesmo que a prova a ser valorada seja a declaração do perito em audiência, isso não anula a necessidade de um exame pericial ser registrado em laudo, até para que as partes possam saber seu conteúdo para se preparar para a audiência (Duce, 2013, p. 151). Para isso, é fundamental que os laudos periciais informem como os exames foram realizados (Vázquez, 2023, p. 03).

O laudo deve detalhar a inferência pericial, isto é, quais foram as premissas maior (generalizações empíricas utilizadas e os seus fundamentos em certas áreas do conhecimento) e menor (aplicação dessas generalizações aos fatos do caso) consideradas e quais são as conclusões que podem ser tomadas a partir disso (Vázquez, 2023, p. 06-08).

Também é necessário que o laudo informe o objeto da perícia; as qualificações teóricas e práticas do(s) perito(s); as pessoas que participaram dos exames e quais tarefas realizaram; os métodos e técnicas empregados, bem como a fiabilidade que lhes é atribuída; e quais foram as informações com as quais o perito teve contato ao realizar seus exames, bem como todos os dados que foram produzidos nessa análise. Em resumo, deve detalhar o raciocínio pericial de uma forma que permita sua reanálise por outros peritos (Vázquez, 2023, p. 27-28).

Se o laudo apenas estabelece conclusões, mas não detalha como se deram as inferências periciais que as embasam, cria um obstáculo para contraditá-lo (Vázquez, 2023, p. 06) e prejudica o controle da qualidade da prova pericial.

No que se refere à limitação de vieses cognitivos, talvez fosse melhor que os sistemas processuais prestassem mais atenção para técnicas de produção da prova pericial mais adequadas para isso.

Para limitar as possibilidades de um perito incorrer em viés de contexto, por exemplo, é recomendável que não lhe seja concedido acesso à íntegra do processo: se o objetivo é não ter contato com informações irrelevantes, é aconselhável que se precise quais são os elementos necessários para a análise (Vázquez, 2023, p. 31; President's Council of Advisors on Science and Technology, 2016, p. 99).

---

[12] A respeito disso, esclarece-se que não se desconhecem os riscos que um contra-exame mal realizado representa para a prova testemunhal. Sobre o tema, ver De Paula Ramos, 2023.

A ordem de apreciação dos elementos pelos peritos também importa. Para não terem o raciocínio contaminado, recomenda-se que primeiro tenham contato com as evidências que estão analisando, registrem suas conclusões e depois tenham contato com o material referenciado, o que pode evitar que o último comprometa o que se entende sobre as evidências (Dror *et al*, 2015, p. 03; Kassin, Dror, Kukucka, 2013, p. 49; President's Council of Advisors on Science and Technology, 2016, p. 65).

O registro das conclusões parciais também deve ser indicado no laudo final, até para demonstrar se houve alguma mudança de posição. A indicação de uma alteração sobre essas conclusões não é capaz de por si só evitar que os peritos incorram em algum viés cognitivo, mas pode torná-los mais transparentes (Dror *et al*, 2015, p. 03) e limitar seus efeitos.

Ademais, caso seja produzida uma perícia posterior sobre uma matéria já periciada, o ideal é que o segundo perito não saiba quais foram as conclusões iniciais, pois, caso contrário, tenderá a ficar enviesado e a confirmá-las (Garrett, 2021, p. 119-120; Edmond *et al*, 2015, p. 02; President's Council of Advisors on Science and Technology, 2016, p. 89-90).

Nesse sentido, cita-se estudo empírico que realizou comparação entre revisões de perícias feitas às cegas (em que o revisor não sabe as conclusões do primeiro exame) ou não. Ao final, constatou-se que a taxa de discordância das conclusões iniciais foi cinco vezes maior em verificações às cegas do que em revisões nas quais o perito sabia dessas conclusões (Mattijssen *et al*, 2020) [13].

Por mais que não possa ser automaticamente generalizada, tal conclusão merece atenção. Afinal, parece consenso que, feita às cegas, a revisão de uma perícia por outro perito pode reduzir erros (Edmond *et al*, 2016, p. 04) e, principalmente, limitar a ocorrência de vieses confirmatórios (Kassin, Dror, Kukucka, 2013, p. 49).

Para inibir o viés de confirmação, também é recomendável que os peritos não façam suas análises à luz de uma única hipótese sobre o que estão examinando (Vázquez, 2023, p. 31). Se isso ocorrer, aumentam-se as chances de ser adotado um raciocínio circular, sendo as evidências analisadas de acordo com essa hipótese.

Essas ideias consistem em boas práticas de produção de provas periciais. Para limitar os efeitos dos vieses cognitivos, os juízes, em suas decisões, deveriam verificar se a produção da perícia observou essas técnicas (President's Council of Advisors on Science and Technology, 2016, p. 101) [14].

---

[13] Quanto menos informações sobre o exame anterior sejam concedidas ao perito que realizará a análise posterior, maiores as chances de que essa cumpra a função de revisão. Nesse mesmo estudo, os pesquisadores também constataram que peritos têm maior tendência a concordar com as conclusões iniciais quando sabem que foi realizado por alguém hierarquicamente superior. No caso dessa pesquisa, a tendência encontrada de concordância com o exame inicial foi 2,5 vezes maior nos casos em que se sabia que o perito inicial tinha uma posição superior, quando comparados com os casos em que era sabido que o examinador anterior ocupava um cargo inferior (Mattijssen *et al*, 2020).

[14] Apesar de essa recomendação ter sido feita em trecho específico sobre perícias de identificação de impressões digitais, não há motivo para restringi-la a esse tipo de exame.

Referidos mecanismos muito podem contribuir na tarefa de mitigar a parcialidade cognitiva e diminuir o número de erros judiciais. Contudo, são pouco empregados pelos sistemas processuais, que continuam adotando instrumentos de capacidade duvidosa.

Por mais que o contraditório seja uma ferramenta poderosa para controlar a atividade probatória, pode ter pouca aplicabilidade se não há atenção às particularidades da prova pericial. Se os mecanismos de contraditório não são capazes de mitigar os efeitos da parcialidade cognitiva — o que não são— a garantia da qualidade desse meio de prova fica ainda mais dificultada.

As ferramentas instituídas para evitar os vieses cognitivos devem ter respaldo empírico e focar em questões normativas e de desenho institucional (Páez, 2021, p. 205-206). Mas, como visto, os mecanismos tradicionais de contradição não têm a capacidade necessária para lidar com questões relativas à parcialidade cognitiva.

Não se pode confiar no contraditório como mecanismo exclusivo para lidar com a garantia da qualidade da prova pericial, pois não serve para limitar a parcialidade cognitiva e seus efeitos.

Para isso, seria muito melhor que fossem empregadas ao menos algumas das técnicas descritas acima. Por serem mais aptas para inibir vieses cognitivos, em muito poderiam colaborar para termos perícias de maior qualidade.

Isso também poderia fortalecer o contraditório sobre a prova pericial, pois o controle da sua qualidade ficaria mais objetivo. Sabendo qual foi o material que o perito teve contato, a parte pode questioná-lo para verificar qual foi a influência de certa informação irrelevante nas suas conclusões; se a revisão às cegas de uma perícia aponta equívocos na primeira perícia, a parte terá maior capacidade de demonstrar, no processo, porque essa deve ser descartada.

Seria fortalecida a efetividade da reação contrária, objetivo da concepção moderna do contraditório. Logo, não deve haver uma oposição entre os mecanismos tradicionais de contradição e as boas práticas de produção de perícias: utilizadas em conjunto, seriam mais efetivas para cumprir o objetivo que partilham, a diminuição do risco de provas periciais serem fontes de erro.

Mas, como será detalhado no próximo capítulo, isso parece distante para o processo penal brasileiro, que confia apenas no contraditório para controlar a qualidade da prova pericial. Sem boas práticas complementares de produção de perícias, o contraditório não apenas falha em mitigar os riscos inerentes aos vieses como pode, em determinadas configurações, criar condições para ampliá-los.

Assim, será analisado se o processo penal brasileiro está estruturado para lidar com o problema da parcialidade cognitiva na produção da prova pericial. Com essa preocupação, o trabalho abordará o caso específico dos institutos dos quesitos e dos quesitos complementares e analisará como podem ser compreendidos em meio a esse objetivo.

3. PRODUÇÃO DA PROVA PERICIAL NO PROCESSO PENAL BRASILEIRO

No processo penal brasileiro, a regra é as provas periciais serem confeccionadas por um perito oficial, que deve ser um servidor público concursado para atuar em um órgão oficial, conforme os artigos 159 do Código de Processo Penal (adiante, CPP) e 2º da Lei n. 12.030/2009 (BRASIL. Lei nº 12.030, de 17 de setembro de 2009. Dispõe sobre as perícias oficiais e dá outras providências.).

Como recorre a instituições oficiais para a produção de perícias, o modelo adotado é o de designação de peritos pelo juiz (Vázquez, 2021, p. 431-432). Tanto que o CPP trata os peritos como auxiliares do juiz.

Assim, o CPP também se preocupa com a imparcialidade dos peritos. Por conta disso, estipula, no artigo 279, situações em que um perito não pode exercer a sua função, dentre as quais está a prevista no inciso II, que consiste no fato de ter prestado depoimento no processo ou opinado sobre o objeto da perícia, o que tem especial relevância para este trabalho.

Por ser um modelo de perito de confiança do juiz, seria esperado que, em respeito ao contraditório, as partes pudessem participar da produção da prova pericial desde seu início. Afinal, nesse modelo, deveria ser permitido às partes acompanhar a formação dessa prova para poderem questioná-la adequadamente (Vázquez, 2021, p. 451). Contudo, isso não ocorre na prática processual penal brasileira, o que influencia toda essa etapa.

3.1. Contexto da produção da prova pericial

Na maioria dos casos brasileiros, as provas periciais são produzidas durante o inquérito policial (Amaral, 2023, p. 68). Como visto, isso ocorre muito por causa da natureza dos elementos examinados, que devem ser apreciados o quanto antes para essa prova ter maior qualidade. Só que isso repercute em todo o procedimento de produção dessa prova.

Diferentemente do ideal, os órgãos oficiais responsáveis pelos exames periciais no Brasil não são instituições completamente autônomas e próximas à comunidade acadêmica. Dos 27 estados brasileiros, em 18 esses institutos periciais estão vinculados às Secretarias de Segurança Pública, enquanto nos outros 9 e na esfera federal fazem parte da polícia judiciária (Medeiros, 2020, p. 11-12)[15].

[15] Mesmo em estados onde o instituto está vinculado à secretaria, isso não impede que a polícia judiciária tenha peritos trabalhando diretamente em algumas divisões suas, como ocorre no Estado de São Paulo (Decreto nº 42.847 do Estado de São Paulo, de 9 de fevereiro de 1998, dispõe sobre a estrutura organizacional da Superintendência da Polícia Técnico-Científica e dá providências correlatas).

De uma forma ou outra, todos estão inseridos na lógica do sistema de combate à criminalidade. Mesmo que não estejam vinculados à polícia judiciária, a inserção dos institutos periciais no organograma das Secretarias de Segurança Pública faz com que os órgãos que investigam crimes e os que realizam perícias estejam sob a mesma subordinação, o que aumenta as dúvidas sobre a imparcialidade dos peritos (Amaral, 2023, p. 68).

Há uma grande associação entre peritos e polícia judiciária. Não à toa, existem definições de peritos oficiais como aqueles que pertencem aos quadros da polícia (Badaró, 2020, p. 498) ou como membros de um órgão público que geralmente é a polícia judiciária (Pacelli, Fischer, 2016, p. 399).

Assim, não surpreende que, em estudo sobre a situação das perícias criminais no Brasil, Flávia Medeiros tenha constatado que esses institutos oficiais funcionam sob os moldes de um corpo policial, como se fossem um órgão de acusação, o que compromete a realização de perícias idôneas (Medeiros, 2020, p. 25)[16].

Como abordado anteriormente, peritos que se entendem como parte do sistema de combate à criminalidade tendem a ficar cognitivamente parciais[17]. Se pensam que têm o papel de fornecer evidências incriminatórias, ficam mais vulneráveis ao viés confirmatório e, consequentemente, a chegarem a conclusões equivocadas. Diante disso, conclui-se que o contexto de produção da prova pericial no processo penal brasileiro expõe os peritos a esse viés.

Agora, passa-se a descrever o procedimento de produção dessa prova para verificar se as regras processuais deixam os peritos mais expostos a isso — e também ao viés de contexto — e se existem instrumentos para tentar limitar esses vieses.

3.2. Rito de produção da prova pericial

Em 2008, o CPP brasileiro passou por reformas em diversos temas, dentre os quais a disciplina legal das provas. O regramento das provas era a matéria de atualização mais necessária, pois instituído em um regime autoritário (Gomes Filho, 2008, p. 247).

---

[16] A identificação desse comportamento em casos concretos é difícil. Um exemplo disso é um caso no qual uma prova pericial só foi anulada pelo Superior Tribunal de Justiça após a perita que a elaborou admitir, em um *podcast*, que forçou a suspeita a confessar (AgRg no HC 898.724/SP, de 12 de dezembro de 2024). Inclusive, nesse caso, a perita integrava a polícia judiciária, o que indica como a vinculação entre peritos e polícia pode gerar exames periciais parciais.

[17] Apesar de envolver como os peritos são originalmente designados, isso não os torna contaminados por parcialidade de origem, sendo uma questão de parcialidade cognitiva. Adota-se, aqui, o entendimento de que a parcialidade de origem dependeria de uma vontade mais forte do perito do que a cognitiva (Vázquez, 2021, p. 132). Assim, a alocação institucional dos órgãos periciais não faria com que os peritos tenham uma vontade consciente de indicar elementos incriminadores, mas sim uma predisposição maior a apontá-los em seus exames.

Nessa reforma, foi feita uma distinção entre prova e elementos informativos: a primeira é produzida mediante contraditório judicial, ao passo que os últimos são colhidos durante a fase de investigação. Logo, o contraditório foi consagrado como o melhor método para a obtenção da verdade (Gomes Filho, 2008, p. 251).

Assim, o artigo 155 do CPP passou a prever que uma sentença não pode ser fundamentada exclusivamente em elementos informativos, mas fez uma exceção para as provas antecipadas, cautelares e irrepetíveis, o que tem grande repercussão em provas periciais.

Afinal, grande parte das provas periciais precisa ser produzida o quanto antes para ter maior qualidade. É esse o motivo que faz a maioria dessas provas ser produzida na fase de investigação, embora o CPP permita que sejam confeccionadas na fase judicial (Badaró, 2020, p. 502).

Na doutrina brasileira, encontram-se caracterizações dessas perícias como provas cautelares (Badaró, 2020, p. 472) ou irrepetíveis (Corrêa, 2006, p. 237). Mas, por uma categorização ou por outra, há concordância sobre a inviabilidade de reproduzir esses exames em fase processual (Badaró, 2020, p. 472; Corrêa, 2006, p. 237).

Não interessa ao trabalho discutir se provas periciais são irrepetíveis ou cautelares, nem debater esses conceitos. A questão é que há consenso sobre essas perícias poderem ser produzidas sem incidência plena do contraditório e mesmo assim serem utilizadas para fundamentar exclusivamente uma sentença.

Nessas situações, o contraditório é exercido em sua modalidade diferida, com a reação das partes feita apenas posteriormente, na fase processual. Como o contraditório diferido não permite que as partes influenciem imediatamente a percepção sobre o que está sendo produzido, a qualidade e a possibilidade das objeções que podem ser feitas é limitada (Pacelli, Fischer, 2016, p. 347; Badaró, 2020, p. 473).

Em linha com a consagração do contraditório como melhor método de formação das provas, a reforma de 2008 estabeleceu uma nova disciplina das perícias, mais adequada a isso (Gomes Filho, 2008, p. 247). Contudo, o CPP não detalha um procedimento para produção desse meio de prova.

No artigo 160, o CPP apenas dispõe que os exames periciais serão relatados em um laudo, que conterá descrição detalhada do que foi analisado e respostas aos quesitos. O laudo pericial é um documento escrito[18] pelo qual os elementos que constituem a prova pericial ingressam ao processo, pois os exames são feitos de forma extraprocessual (Gomes Filho, 2022b, p. 473).

[18] Portanto, no processo penal brasileiro, a prova pericial também possui uma dimensão documental. Como muitas vezes esses laudos possuem imagens ou outros tipos de representações visuais do que foi analisado, deve se ter em mente que esses sinais indexicais captam apenas um ângulo ou momento do objeto examinado (De Paula Ramos, 2025, p. 188), o que deve ser considerado na valoração dessa prova.

Também não há uma determinação sobre a forma desse laudo. Idealmente, seria composto por um preâmbulo, em que é indicado o perito que o elaborou, a natureza da perícia e o objeto analisado; um histórico dos exames realizados; uma descrição do que foi constatado; uma discussão sobre os diagnósticos do perito, com a sua respectiva fundamentação empírica; as conclusões tomadas; e, ao final, as respostas aos quesitos (Gomes Filho, 2022b, p. 473)[19]. Mas, sem uma exigência legal, facilita-se que alguma dessas partes seja suprimida[20].

Por sua vez, os quesitos são questões formuladas aos peritos sobre o assunto que examinarão (Nucci, 2024, p. 361). Consistem na forma de comunicação entre quem solicita a prova pericial e quem a produz. Todavia, não há nenhuma disposição legal sobre como devem ser elaborados.

Na maioria dos casos, os quesitos encaminhados são aqueles previstos em formulários sobre cada tipo de perícia (Badaró, 2020, p. 502). A Polícia Federal, por exemplo, possui um manual de orientação de quesitos, em que são indicados quais devem ser feitos e quais não devem ser (Departamento de Polícia Federal, 2012).

Ademais, o CPP não indica como essas respostas devem ser redigidas. Na doutrina, é possível encontrar uma recomendação por respostas que sejam claras e diretas, preferencialmente com a utilização de sim ou não (Gomes Filho, 2022b, p. 473).

Também não há qualquer determinação sobre quais elementos devem ser e quais não podem ser fornecidos aos peritos. É perfeitamente possível que os autos sejam integralmente encaminhados ao perito, que pode ter contato com informações irrelevantes antes de realizar a sua análise e, portanto, ser contaminado por um viés contextual.

Sem uma exigência legal sobre o conteúdo do laudo, não é raro que não se saiba quais foram os elementos com os quais o perito teve contato, quais as generalizações empíricas consideradas ou quais métodos e técnicas foram empregados nos exames, o que prejudica a tarefa de contraditá-lo.

Há também um problema causado por uma redação legal inadequada. É fato que o artigo 176 do CPP permite que as partes e a autoridade judicial formulem quesitos antes de realizada a perícia. Mas como só existem partes na fase processual, prevalece que esse dispositivo não tem aplicabilidade para provas produzidas na fase

---

[19] Em sentido similar, Gustavo Badaró afirma que o laudo é composto por quatro partes: preâmbulo, exposição, discussão e conclusão (Badaró, 2020, p. 503).

[20] Não raras vezes, laudos periciais não informam qual foi o método adotado. Para superar esse problema, e dada a lacuna do CPP, seria ideal que fossem empregados, ao processo penal, os artigos 473, inciso III, e 479 do Código de Processo Civil, que dispõem que o laudo pericial deverá indicar o «método utilizado, esclarecendo-o e demonstrando ser predominantemente aceito pelos especialistas da área do conhecimento da qual se originou» e que, ao valorar uma prova pericial, o juiz deve informar «os motivos que o levaram a considerar ou a deixar de considerar as conclusões do laudo, levando em conta o método utilizado pelo perito», respectivamente (Kircher, 2019, p. 160).

de inquérito policial (Gomes Filho, 2008, p. 277), na qual a maioria costuma ser confeccionada.

Portanto, em perícias produzidas durante o inquérito policial, apenas costumam ser formulados quesitos pela autoridade policial ou pelo Ministério Público, sem participação da defesa (Corrêa, 2006, p. 237; Pacelli, Fischer, 2016, p. 405). Essa situação ocorre na maioria dos casos, independente da natureza da perícia.

Sem que essa prova seja produzida mediante contraditório desde o início, aumentam-se as chances de ocorrerem equívocos (Nucci, 2024, p. 414). Esse cenário fica ainda mais grave quando ocorre em exames periciais que não têm como serem reproduzidos nas mesmas condições na fase processual (Corrêa, 2006, p. 237), que seria o primeiro momento em que a defesa poderá formular quesitos (Pacelli, Fischer, 2016, p. 405).

Nessa fase, segundo o artigo 159, §§ 3º e 5º, do CPP, a defesa pode fazer uso de três ferramentas de contraditório: requerer que o perito compareça em audiência para esclarecer o laudo; formular quesitos para complementação da perícia; e indicar um assistente técnico, que poderá apresentar um parecer sobre a perícia oficial e ser inquirido em audiência.

A instituição dessas três ferramentas é tida como uma inovação positiva da reforma de 2008, pois fortalece o contraditório na prova pericial (Gomes Filho, 2008, p. 275), aprimorando as reduzidas possibilidades existentes no regime diferido (Pacelli, Fischer, 2016, p. 349-350).

Assim, na prática, a prova pericial costuma ser produzida durante o inquérito policial por um perito oficial com quesitos formulados apenas pela autoridade policial; após a juntada do laudo aos autos, já na fase processual, as partes poderão fazer uso de um dos três instrumentos referidos acima (Badaró, 2020, p. 472-473). É esse procedimento que podemos chamar de rito de produção da prova pericial no processo penal brasileiro[21].

Nesse rito, a defesa apenas tem os três mecanismos de contraditório citados acima para tentar controlar a qualidade da prova pericial. Contudo, conforme visto, o contraditório, sem ser acompanhado por boas práticas de produção das perícias, não é a melhor ferramenta para lidar com a parcialidade cognitiva.

Como não há a previsão legal de nenhuma das boas práticas para evitar parcialidade cognitiva na produção das provas periciais, é possível sustentar que o processo penal brasileiro não está adequadamente estruturado para lidar com o problema dos

[21] A dificuldade da participação defensiva na fase de investigação decorre de uma interpretação literal dos dispositivos legais sobre a prova pericial. Todavia, em recente julgado, o Superior Tribunal de Justiça reconheceu o direito de investigado de constituir assistente técnico para atuar na produção da prova pericial já na fase de inquérito, o que pode ser um sinal de que esse quadro está sendo alterado (RHC 200.979/SP, de 15 de dezembro de 2025).

vieses cognitivos, pois confia exclusivamente no contraditório para controlar a qualidade das perícias.

Mais do que isso: como antecipado no capítulo anterior, um rito que deposita no contraditório toda a responsabilidade pelo controle da qualidade das perícias pode, em vez de mitigar os vieses cognitivos, criar condições institucionais para ampliá-los. É o que o estudo dos quesitos e dos quesitos complementares permitirá demonstrar.

3.2. Caso de estudo: quesitos e quesitos complementares

Dentre esses instrumentos de contraditório, este trabalho focará apenas na formulação de quesitos complementares. A escolha desse mecanismo se deve a ser o mais parecido com a prerrogativa que a autoridade que solicita essa prova tem de formular quesitos ao perito sobre o objeto examinado[22].

Assim, será feito um estudo específico para analisar relações que podem existir entre os quesitos, iniciais e complementares, e os vieses de confirmação e de contexto.

A dinâmica dos quesitos e quesitos complementares tem grande semelhança com a que existe para a prova testemunhal. A parte ou autoridade que solicita a produção da prova primeiro questiona o terceiro que fornecerá conhecimento ao processo. Depois, a parte contrária complementa essa prova com as suas questões.

A diferença é que, pelo rito da prova pericial na prática processual penal brasileira, os quesitos da parte contrária não são formulados no mesmo ato em que feitos os primeiros, mas apenas após a finalização da primeira versão do laudo. Assim, o trabalho examinará se isso pode ter algum impacto na ocorrência de vieses confirmatórios.

Sobre os quesitos complementares, trata-se de uma regra processual compreendida como uma ferramenta de controle da qualidade da prova pericial, que permitiria um amplo esclarecimento sobre o objeto examinado (Gomes Filho, 2008, p. 277). Não causam a produção de uma prova nova, mas sim uma revisão da prova já confeccionada.

Nada impede que o mesmo perito que tenha feito o exame inicial também realize o exame complementar. Afinal, o artigo 279, inciso II, do CPP somente impede que um perito que tenha emitido uma opinião técnica sobre determinado objeto produza uma prova nova sobre esse mesmo material (Pacelli, Fischer, 2016, p. 610)[23].

---

[22] Essa escolha também foi feita por ser o mecanismo de mais fácil utilização pela defesa, pois não é todo acusado que tem condições financeiras de contratar um assistente técnico. Quanto ao comparecimento do perito em juízo, trata-se de instituto que encontra obstáculo para utilização no próprio CPP: por mais que o artigo 159, §5º, inciso I, inicialmente permita que as partes solicitem o comparecimento em juízo do perito para esclarecer o exame e responder a questões complementares, a parte final desse dispositivo permite que, caso queira, o perito formule suas respostas em laudo pericial complementar, sem necessidade de comparecer em juízo.

[23] Não são muitos os casos em que esse dispositivo legal tem seu significado interpretado pelos tribunais, o que torna difícil falar sobre uma posição dominante sobre essa matéria. De toda forma,

Também não há regulação legal sobre o material encaminhado ao perito que responderá os quesitos complementares. Essa lacuna pode ser ainda pior do que a que existe sobre os exames iniciais: mesmo que os exames complementares sejam conduzidos por um perito diferente, nada o impede de ter contato com as conclusões do laudo inicial, sendo contaminado por isso.

Diante disso, questiona-se se esse rito pode contribuir para a ocorrência de vieses cognitivos, o que será analisado no capítulo a seguir.

4. QUESITOS, QUESITOS COMPLEMENTARES E PROBLEMAS RELATIVOS AOS VIESES COGNITIVOS NA PRODUÇÃO DA PROVA PERICIAL NO PROCESSO PENAL BRASILEIRO

No processo penal brasileiro, os quesitos são a forma de comunicação entre quem solicita uma perícia com o perito que irá realizá-la e consistem em perguntas formuladas pela primeira ao segundo.

A respeito disso, esclarece-se que este trabalho não tem como objetivo fazer uma crítica total aos institutos dos quesitos e dos quesitos complementares. Idealmente, teriam grande utilidade para termos perícias de maior qualidade: por serem perguntas sobre o que deve ser explicado, auxiliam a estabelecer o objeto dos exames, o que pode evitar provas periciais com escopo demasiadamente amplo (Vázquez, 2023, p. 30)[24].

O que será analisado é se, dado o rito de produção de perícias no processo penal brasileiro, os quesitos e os quesitos complementares representam um risco para a contaminação dos peritos por vieses de confirmação e de contexto.

4.1. Quesitos no inquérito policial: possível fonte de vieses de confirmação e de contexto

No processo penal brasileiro, a maioria das provas periciais é produzida durante a fase de inquérito policial, sem a participação da defesa ou do juiz. Apenas são

---

é possível citar um caso no qual o Tribunal de Justiça do Estado de São Paulo determinou que essa regra de impedimento «não alcança «experts» que atuam desde o início do feito e prestam meros esclarecimentos tão somete a respeito do objeto da perícia» (APL 0251309-33.2010.8.26.0000, de 1º de junho de 2011). Nesse caso, os esclarecimentos solicitados consistiam justamente em respostas a quesitos apresentados sobre um laudo pericial já finalizado.

[24] Os quesitos e os quesitos complementares também poderiam auxiliar a determinar a relevância da prova pericial, o que seria de grande auxílio para um juízo de admissibilidade (Vázquez, 2023, p. 30). Contudo, como não há um juízo de admissibilidade legalmente previsto para as provas cautelares e irrepetíveis no processo penal brasileiro, no que se enquadram muitas provas periciais, essa potencialidade acaba sendo inutilizada, de forma que não há um filtro epistemológico sobre essas provas (Vieira, 2023, p. 257). Logo, seria recomendável uma alteração legislativa para instituir esse juízo (Vieira, 2023, p. 259-261).

enviados ao perito quesitos formulados pela autoridade policial ou pelo Ministério Público.

Não há nenhum controle sobre a redação desses quesitos. O CPP não estipula nenhuma regra sobre a formulação dessas questões e não há nenhuma autoridade que as aprecie para determinar se contém algum risco de contaminar os peritos. Os quesitos são diretamente encaminhados aos peritos assim como formulados por quem solicita a perícia.

Geralmente, os exames são realizados por um perito oficial, que atua nas instituições estatais responsáveis pelos exames periciais. Referidos órgãos estão inseridos na lógica do sistema de repressão à criminalidade, não sendo institutos completamente autônomos e mais ligados às comunidades científica e acadêmica. Não à toa, costumam adotar comportamentos típicos de um corpo policial.

Esse é o cenário em que a maioria dos exames periciais é feita no processo penal brasileiro. Por tudo exposto anteriormente, conclui-se que é um desenho institucional que aumenta o risco de os peritos serem contaminados por vieses confirmatórios ao realizarem suas análises.

Inseridos no sistema de combate à criminalidade, os peritos têm maiores chances de entenderem que têm o papel de fornecerem evidências incriminatórias. Assim, aumentam-se os riscos de que conduzam exames a partir da expectativa de cumprir uma função que não deveriam ter, sendo contaminados por um viés confirmatório e adotando um raciocínio circular.

Isso tudo pode ser potencializado pela forma como os quesitos são formulados. Como visto, a forma como as perguntas são feitas aos peritos tem o perigo de enviesar a análise que será realizada (Amaral, Bruni, 2023, p. 893-894). Seria necessário que o processo penal brasileiro tivesse ferramentas para identificar previamente quesitos com potencial de enviesamento e impedir que sejam encaminhados.

Mas não é isso que ocorre. Os quesitos são encaminhados diretamente pela autoridade que solicita a perícia para quem irá realizar os exames, sem nenhum controle sobre a redação dessas perguntas.

É completamente possível que sejam formulados quesitos com perguntas sugestivas aos peritos. Tratando especificamente da prova testemunhal, Vitor de Paula Ramos conceitua essas perguntas como «questões que visam a dirigir a testemunha para confirmar aquilo que o entrevistador pretende, consciente ou inconscientemente» (De Paula Ramos, 2023, p. 199) [25].

[25] Embora elaborada em obra que trata da prova testemunhal, não há problema em utilizar essa definição, bem como outras questões relativas à interação entre falante e entrevistador, para temas relativos às provas periciais. Afinal, a prova pericial faz parte do gênero dos testemunhos. Tanto que, logo antes de elaborar esse conceito, o autor cita uma série de estudos sobre o viés confirmatório em provas periciais (De Paula Ramos, 2023, p. 198-199).

Sobre isso, rememora-se que, na maioria das perícias criminais no Brasil, os quesitos são formulados pela autoridade policial que conduz a investigação e, portanto, deveria ter como um dos seus objetivos a busca pela verdade. Por causa disso, é comum que policiais formulem hipóteses sobre o caso ao participarem da investigação (De Paula Ramos, 2023, p. 195-197)[26].

Assim, há um risco de que orientem os seus trabalhos com base na hipótese que têm sobre o caso com o intuito, mesmo que inconsciente, de corroborar o que pensam, sendo influenciados por um viés confirmatório (De Paula Ramos, 2023, p. 198).

Se isso acontecer, os policiais tenderão a fazer perguntas que expressem essa hipótese, o que pode contaminar significativamente as respostas dadas (De Paula Ramos, 2023, p. 199). Se há um suspeito principal, por exemplo, é capaz que façam perguntas que direcionem a resposta a essa pessoa.

Logo, em matéria de perícias, o viés da autoridade policial pode ser transmitido ao perito, o que pode comprometer o resultado dos exames. Ocorreria o efeito bola de neve dos vieses cognitivos, em que uma pessoa enviesada contamina outra. No contexto brasileiro, trata-se de risco real, pois a maioria das perícias é feita apenas com quesitos elaborados pela polícia.

Isso cria outro risco para a incidência de vieses confirmatórios em provas periciais. Se os quesitos da polícia tendem a veicular uma hipótese sobre o caso e são os únicos que costumam ser encaminhados aos peritos, é possível que as análises periciais sejam feitas a fim de confirmar essa hipótese única, ainda que inconscientemente[27].

Portanto, a forma como os quesitos são tratados no processo penal brasileiro aumenta os riscos de os peritos serem contaminados por viés confirmatório. Não só pode haver contaminação porque não há nenhuma proibição de realização de quesitos sugestivos, como também porque os peritos costumam responder apenas quesitos formulados por autoridades que já possuem certo risco de estarem enviesadas.

Soma-se a isso que o próprio desenho institucional dos órgãos periciais expõe os peritos a maiores riscos de contaminação por viés confirmatório, pois estão inseridos na lógica do sistema de combate à criminalidade.

Sobre a ausência de impedimento de quesitos sugestivos, chama atenção que isso não ocorra em relação à prova testemunhal. Em seu artigo 212, o CPP determina que o juiz deve indeferir perguntas que possam induzir a resposta da testemunha.

---

[26] Ressalta-se que, no trecho citado, o autor não está falando apenas de policiais, mas sim sobre pessoas que conduzem investigações, inclusive em diferentes campos do conhecimento. Mas, considerando que os policiais têm como função conduzir investigações, entende-se que não há problemas em utilizar tal formulação neste trabalho.

[27] O problema envolvendo quesitos e hipóteses não é propriamente que esses veiculem hipóteses, mas sim uma hipótese única. Se o início de investigações sobre delitos deve ser orientado pelo raciocínio abdutivo (Moscatelli, 2023, p. 138), é ideal que os quesitos sejam formulados para testar a hipótese que está sendo investigada, mas isso não pode ser uma desculpa para que outras hipóteses possíveis não sejam testadas nesse momento.

Não se trata de defender que devam ser aplicadas à prova pericial todas as previsões existentes para a prova testemunhal[28]. Afinal, os peritos não comparecem ao processo para fazer um exercício de memória como uma testemunha comum. Questões sugestivas, portanto, não têm a capacidade de confundir a memória de peritos e fazer com que imediatamente forneçam respostas equivocadas.

Mas isso não significa que não possam contaminar os peritos. Justamente por delimitar o objeto da perícia, os quesitos delineiam toda essa prova. Redigidos de forma sugestiva, podem criar expectativas nos peritos sobre a investigação que será feita para respondê-los, as operações periciais que serão executadas e até sobre as conclusões tomadas ao final. Se todo esse processo for influenciado por um viés confirmatório, haverá grande chance de a prova pericial ter um resultado errôneo.

Dado o risco que representam, não se pode aceitar que os quesitos sejam encaminhados diretamente aos peritos sem nenhum mecanismo de controle. Também não pode ser aceito que não haja nenhuma regulamentação sobre a forma de redação desses quesitos, que costuma ser delegada a formulários institucionais.

Por mais que esses formulários se assemelhem à ideia de inibir vieses cognitivos por meio do estabelecimento de protocolos (Páez, 2021, p. 25-28), não é possível se contentar com qualquer protocolo. Até porque, caso um formulário encampe uma ideia que aumente os riscos de acontecerem vieses cognitivos, essa finalidade não será atingida[29].

Nesse sentido, no manual de quesitos elaborado pela Polícia Federal, são recomendados quesitos que podem levar peritos a incorrerem em viés confirmatório[30]. Por exemplo, em perícias de reconhecimento de impressões digitais, recomenda-se que seja perguntado se «A impressão digital registrada no documento é a mesma presente na individual datiloscópica do suspeito FULANO DE TAL?» (Departamento de Polícia Federal, 2012, p. 115).26[31]

Trata-se de recomendação que vai na contramão do que se sabe sobre esse tipo de perícia. Conforme estudo empírico realizado por Itiel Dror, David Charlton e Ailsa Péron, a identificação do suspeito é um fator que aumenta o risco de a perícia de

[28] Até porque, um dos maiores problemas para o contraditório da prova pericial é que o seu rito de produção costuma ser o mesmo da prova testemunhal (Salaverría, 2021, p. 187-188).

[29] Ressalta-se que Andrés Páez não defende que qualquer protocolo seria suficiente para limitar vieses cognitivos, pois também sustenta que o combate a esses vieses deve ser feito por práticas que tenham respaldo empírico (Páez, 2021, p. 205-206).

[30] A escolha pela análise do manual da Polícia Federal, e não de outra polícia judiciária, deu-se em razão de essa ser a única com atuação em todos os entes federativos brasileiros.

[31] Quesitos semelhantes também são recomendados outros tipos de perícias, como a de reconhecimento de indivíduos por imagens, para as quais inclusive é recomendado que não sejam formulados quesitos que não indiquem nominalmente o suspeito (Departamento de Polícia Federal, 2012, p. 26-27).

identificação de impressões digitais ser cognitivamente contaminada (Dror, Charlton, Péron, p. 2006)[32].

Essa situação também ilustra como quesitos mal redigidos podem causar um viés de contexto nos peritos. Afinal, não há nenhum impeditivo para que informações irrelevantes — como o nome do suspeito, no exemplo citado — sejam veiculadas nesses quesitos.

Ao introduzir uma informação irrelevante, a própria pergunta que delimita o trabalho do perito pode orientar a percepção dele antes de iniciados os exames e comprometer os resultados. Se mal redigidos, os quesitos são uma fonte potencial dos dois vieses analisados neste trabalho.

Portanto, por mais que possa haver boas intenções na formulação de quesitos — o próprio manual da Polícia Federal possui boas recomendações[33] —, a redação dessas perguntas pode enviesar quem as responderá. É preciso maior cuidado com essa atividade.

4.2 Quesitos complementares: possível fonte de viés de confirmação

Uma vez que a maioria das provas periciais no processo penal brasileiro é produzida na fase de inquérito policial, não costuma ser permitido que a defesa formule quesitos antes da realização dos exames, conforme explicado anteriormente. Isso apenas poderá ser feito na fase de ação penal, por meio dos quesitos complementares.

Tal qual ocorre com os quesitos iniciais, não há nenhum controle ou regulamentação sobre a redação dos quesitos complementares. Logo, os riscos de que ocorram vieses cognitivos que foram apontados no item anterior também existem para os quesitos complementares.

---

[32] Nesse estudo, os pesquisadores forneceram a cinco especialistas em identificação de impressões digitais diferentes pares de impressões que, no passado, eles já haviam concluído que pertenciam à mesma pessoa. Contudo, adicionaram que esses pares pertenceriam ao caso em que o FBI erroneamente identificou Brandon Mayfield como o responsável por um atentado em Madri. Feitos novos exames, três afirmaram que as impressões não pertenciam à mesma pessoa e outro deu uma resposta inconclusiva; apenas um manteve a sua conclusão original. Mesmo que se trate de uma sugestão forte, conforme os próprios pesquisadores admitem, isso demonstra como saber a pessoa cuja impressão digital está sendo comparada pode comprometer esse exame pericial. Nessa linha, ao fazer uma revisão sobre o caso, o Departamento de Justiça dos Estados Unidos da América afirmou que um dos fatores que causaram o erro foi que, na análise, já se sabia que as impressões comparadas pertenciam a Mayfield, o que fez com que fosse adotado um raciocínio circular que confirmou a identidade das impressões (United States Department of Justice, Office of the Inspector General, 2006, p. 07).

[33] O manual recomenda que não sejam formulados quesitos inespecíficos, sobre o enquadramento legal de fatos ou sobre temas que competem a outras autoridades administrativas ou que demandem análises subjetivas (Departamento de Polícia Federal, 2012). A última dessas preocupações, inclusive, pode auxiliar a evitar vieses cognitivos, pois, como visto, as chances de que ocorram são maiores quando há maior espaço para subjetividade.

No entanto, o rito de produção da prova pericial no processo penal brasileiro representa um risco ainda maior para a incidência de um viés de confirmação nas respostas aos quesitos complementares.

Como os quesitos complementares só são ser formulados após a finalização do laudo pericial, já terão sido tomadas conclusões iniciais sobre os elementos periciados quando as respostas aos complementares forem elaboradas. É aqui que há o risco de enviesamento.

Isso porque não há nenhuma determinação sobre o material que será encaminhado junto aos quesitos complementares. Nesse cenário, não há nada que impeça o perito que responderá esses quesitos de ter contato com o laudo inicial.

Como visto, na revisão de uma perícia, o ideal é que o revisor não tenha contato com as conclusões iniciais, pois poderá ficar enviesado e apenas confirmá-las. Na revisão via quesitos complementares, o ideal é que o perito revisor somente saiba as premissas e técnicas empregadas pelo primeiro perito, a fim de examinar se possuem respaldo — o que torna ainda maior a necessidade de essas informações constarem no laudo pericial—, e quais conclusões podem ser adotadas a partir disso. Assim, não deve ter contato com as interpretações do primeiro perito (Mattijssen *et al*, 2020)

Há um consenso de que a revisão de perícias deve ser feita sem contato com as conclusões iniciais, o que aumenta as chances de evitar erros e enviesamentos. É inclusive uma técnica de limitação de viés confirmatório que possui respaldo empírico (Mattijssen *et al*, 2020), mas que não é adotada na produção de perícias criminais no Brasil.

Assim, é possível que o perito que responderá os quesitos complementares saiba das conclusões iniciais ao realizar os seus exames, o que aumenta as chances de formular respostas que apenas confirmem o que já foi afirmado antes.

A ideia de revisão do laudo pericial pelos quesitos complementares também deve ser questionada, uma vez que não há nenhum impeditivo para que o perito que responda os quesitos complementares seja o mesmo que realizou o laudo inicial.

Isso conflita com o próprio objetivo desse instituto, pensado para melhorar as possibilidades de controle da qualidade prova pericial ao permitir um amplo esclarecimento sobre o objeto.

Se já existem indícios empíricos de que o contato do perito revisor com as conclusões iniciais pode causar um viés de confirmação, não é nenhum exagero pensar que esse risco será ainda maior se não houver uma distinção entre perito inicial e revisor. Se ambas as funções forem desempenhadas pela mesma pessoa, ela terá todos os incentivos para, ao apreciar os quesitos complementares, chegar a respostas que apenas corroborem suas conclusões iniciais.

Assim, os quesitos complementares são um instituto que visa melhorar a qualidade da prova pericial, mas que, dada a fase em que previstos no processo penal brasileiro, podem apenas aumentar as chances de que ocorram vieses confirmatórios. Se não

há preocupação sobre os quesitos iniciais aumentarem as chances de os peritos serem cognitivamente comprometidos e sobre como serão respondidos os complementares, esses podem ser inúteis para melhorar a qualidade da perícia.

Sem uma atenção a isso, os quesitos complementares não passam de uma ilusão de contraditório, que trabalharão com material que tem chances de ter sido fruto de uma análise enviesada e cujo resultado que atingirá pode ser justamente o contrário do que foi estruturado para fazer.

Isso ilustra como não podemos confiar apenas no contraditório para tratar da prova pericial, o que acontece no processo penal brasileiro. Por não ser a ferramenta mais habilitada para lidar com o problema dos vieses cognitivos, pode apenas aumentar os riscos de que esses vieses ocorram, como acontece com os quesitos complementares.

Para lidar com esse problema, o ideal seria adotarmos as boas práticas de produção da prova pericial descritas no primeiro capítulo. Por exemplo, ao invés de estabelecer os quesitos complementares como um mecanismo de contradição, seria melhor que existisse um sistema de revisão às cegas de perícias. Isso seria melhor para evitar a ocorrência do viés confirmatório e, consequentemente, a realização de perícias equivocadas.

5. CONCLUSÃO

É possível concluir que o processo penal brasileiro não está estruturado para lidar com os vieses de confirmação e de contexto na prova pericial. As perícias são tratadas apenas por regras legais, dispostas em um rito que confia que o contraditório, única e exclusivamente, garantirá a qualidade dessa prova.

Contudo, o contraditório não é a melhor ferramenta para tratar desses vieses cognitivos, isto é, impedir que aconteçam e limitar os seus efeitos. Para tanto, seria melhor se confiássemos em boas práticas de produção de perícias — que poderiam até fortalecer a aplicação do contraditório.

Com um desenho normativo mal concebido, é possível que mecanismos de contraditório apenas incentivem a ocorrência desses vieses. O tema dos quesitos e dos quesitos complementares é um bom exemplo. Da forma em que concebida, sem muita preocupação com eventual comprometimento cognitivo, a comunicação entre quem solicita uma perícia e o perito importa sério risco de que o último seja enviesado.

No que se refere exclusivamente aos quesitos complementares, o risco é de que, sem uma regulamentação sobre o que será encaminhado ao perito revisor e sobre quem pode exercer esse papel, essa revisão seja feita de forma enviesada para corroborar o que foi concluído no laudo pericial inicial.

Mais do que a confiança exclusiva em instrumentos de contradição, seria necessário que o processo penal brasileiro adotasse as boas práticas de produção de prova pericial.

É preciso que exista algum controle sobre a formulação dos quesitos antes de que cheguem à apreciação dos peritos. Isso evitaria que os quesitos veiculem apenas uma hipótese única, bem como que sejam redigidos de forma sugestiva ou que veicule informações desnecessárias.

No contexto organizacional em que estão inseridos os institutos periciais brasileiros, que já aumenta a probabilidade de peritos serem contaminados por viés confirmatório, essa medida poderia mitigar as chances de que esse viés ocorra. Também pode diminuir a probabilidade de contaminação pelo viés contextual, pois teria a capacidade de impedir que os quesitos contenham informações desnecessárias para a atividade pericial.

Por fim, para que os quesitos complementares de fato sejam um instituto que aprimore a qualidade das provas periciais, é necessário que sejam respondidos por um perito diferente do que elaborou o laudo inicial e que não saiba as conclusões iniciais, para não incorrer em um viés confirmatório no seu exame.

Com a adoção de algumas dessas práticas, estaríamos mais perto de termos provas periciais de boa qualidade e, portanto, evitarmos erros judiciais.

BIBLIOGRAFIA

Amaral, Maria Eduarda Azambuja. (2023). *Entre a ciência forense e o processo penal: um modelo interdisciplinar da cadeia de custódia*. D'Plácido.

Amaral, Maria Eduarda Azambuja; Bruni, Alice Thais. (2023). Prova Pericial no Processo Penal: a compreensão e a mitigação dos erros forenses como mecanismo de respeito ao contraditório, à ampla defesa e ao direito à prova lícita. *Revista Brasileira de Direito Processual Penal*, *9*(2), 877-912.

Amaya, Amalia (2022). Virtue and objectivy in legal reasoning. Em G. Villa-Rosas; J. L. Fabra-Zamora. *Objectivity in Jurisprudence, Legal Interpretation and Practical Reasoning*. (271-292). Edward Elgar Publishing.

Badaró, Gustavo Henrique. (2019). *Epistemologia judiciária e prova penal*. Revista dos Tribunais.

Badaró, Gustavo Henrique. (2020). *Processo penal*. (8ª ed.). Revista dos Tribunais.

Cooper, Glinda. S.; Meterko, Vanessa. (2019). Cognitive bias research in forensic science: a systematic review. *Forensic Science International*, *297*, 35-46.

Corrêa, Cristiane da Rocha. (2006). O princípio do contraditório e as provas irrepetíveis no inquérito policial. *Revista Brasileira de Ciências Criminais*, *14*(60), 223-253.

De Paula Ramos, Vitor. (2023). *Prova Testemunhal – Do Subjetivismo ao Objetivismo, do Isolamento Científico ao Diálogo com a Psicologia e a Epistemologia*. (4ª ed.). JusPodivm.

De Paula Ramos, Vitor. (2025). *Prova Documental – Do Documento aos Documentos – Do Suporte à Informação*. (4ª ed.). JusPodivm.

Departamento de Polícia Federal. *Manual de orientação de quesitos da perícia criminal*. (1ª ed). Diretoria Técnico Científica, 2012.

Dror, Itiel. E. (2020). Cognitive and Human Factors in Expert Decision Making: six fallacies and the eight sources of bias. *Analytical chemistry*, *92*(12), 7998–8004.

Dror, Itiel E.; Charlton, David; Péron, Ailsa. E. (2006). Contextual information renders experts vulnerable to making erroneous identifications. *Forensic science international*, *156*(1), 74-78.

Dror, Itiel E.; Cole, Simon A. (2010). The vision in "blind" justice: expert perception, judgment, and visual cognition in forensic pattern recognition. *Psychonomic Bulletin & Review*, *17*(2), 161-167.

Dror, Itiel E.; Thompson, William. C.; Meissner, Christian A.; Kornfield, Irv; Krane, Dan; Saks, Michael; Risinger, Michael (2015). Context Management Toolbox: a linear sequential unmasking (LSU) approach for minimizing cognitive bias in forensic decision making. *Journal of forensic sciences*, *60*(4), 1-5.

Duce, Mauricio. (2014). Derecho a confrontación y uso de declaraciones emitidas en un juicio previo anulado. *Política Criminal*, *9*(17), 118-146.

Duce, Mauricio. (2013). *La prueba pericial: aspectos legales y estratégicos claves para el litigio en los sistemas procesales penales acusatorios*. Didot.

Duce, Mauricio. (2022). Los errores, la aplicación al caso y los sesgos cognitivos de los peritos. Em C. Vázquez. (Coord.). *Manual de Prueba Pericial* (143-184). Suprema Corte de Justicia de la Nación.

Edmond, Gary.; Tangen, Jason; Searston, Rachel; Dror, Itiel E. (2015). Contextual bias and cross-contamination in the forensic sciences: the corrosive implications for investigations, plea bargains, trials and appeals. *Law, Probability and Risk*, *14*(1), 1–25.

Edmond, Gary; Found, Bryan; Martire, Kristy; Ballantyne, Kaye; Hamer, David; Searston, Rachel; Thompson, Matthew; Cunliffe, Emma; Kemp, Richard; San Roque, Mehera; Tangen, Jason; Dioso-Villa, Rachel; Ligertwood, Andrew; Hibbert, David; White, David; Ribeiro, Gianni; Porter, Glenn; Towler, Alice; Roberts, Andrew (2016). Model forensic science. *Australian Journal of Forensic Sciences*, *48*(5), 496–537.

Edmond, Gary. (2020). Regulating forensic science and medicine evidence at trial: It's time for a wall, a gate and some gatekeeping. *Australian Law Journal*, *94*, 427–438.

Ferrer Beltrán, Jordi. (2017). *Prova e verdade no direito*. Revista dos Tribunais.

Ferrer Beltrán, Jordi. (2023). *Valoração racional da prova*. (3ª ed.). JusPodivm. Garrett, Brandon (2021). *Autopsy of a crime lab: exposing the flaws in forensics*. University of California Press.

Garrett, Brandon. (2011). *Convicting the innocent: where criminal prosecutions go wrong*. Harvard University Press.

Gascón Abellán, Marina. (2013). Prueba Científica: un mapa de retos. Em C. Vázquez (Ed.). *Estándares de prueba y prueba científica: ensayos de epistemologia jurídica* (188-201). Marcial Pons.

Giannelli, Paul C. (2010). Independent crime laboratories: the problem of motivational and cognitive bias. *Utah Law Review*, *2*, 247-266.

Gomes Filho, Antonio Magalhães. Provas: Lei 11.690. de 9/6/2008. (2008). Em M. T. R. A. Moura (Coord.). *As reformas no processo penal: as novas leis de 2008 e os projetos de reforma* (246-297). Revista dos Tribunais.

Gomes Filho, Antonio Magalhães. Art. 159. (2022a). Em A. M. Gomes Filho; A. Z. Toron; G. H. Badaró, (Coords.). *Código de Processo Penal comentado* (5ª ed., 469-472). Revista dos Tribunais.

Gomes Filho, Antonio Magalhães. Art. 160. (2022b). Em A. M. Gomes Filho; A. Z. Toron; G. H. Badaró, (Coords.). *Código de Processo Penal comentado* (5ª ed., 473). Revista dos Tribunais.

Kassin, Saul M.; Dror, Itiel E.; Kukucka, Jeff. (2013). The forensic confirmation bias: problems, perspectives, and proposed solutions. *Journal of Applied Research in Memory and Cognition*, *2*(1), 42-52.

Kircher, Luís Felipe Schneider. (2019). A valoração da prova pericial no processo penal brasileiro. *Revista Brasileira de Ciências Criminais*, *27*(153), 133-169.

Malan, Diogo. (2015). Processo Penal aplicado à criminalidade econômico-financeira. *Revista Brasileira de Ciências Criminais*, *114*, 279-320.

Mattijssen, Erwin J. A. R.; Witteman, Cilia L. M.; Berger, Charles E. H.; Stoel, Reinoud D. (2020). Cognitive biases in the peer review of bullet and cartridge case comparison casework: a field study. *Science & Justice*, *60*, 337-346.

Medeiros, Flavia. (2020). *Políticas de Perícia Criminal na Garantia dos Direitos Humanos: relatório sobre a autonomia da perícia técnico-científica no Brasil*. Friedrich-Ebert-Stiftung (FES).

Moscatelli, Livia. (2023). La importancia de la abducción en la etapa de investigación criminal. *Quaestio Facti. Revista Internacional Sobre Razonamiento Probatorio*, *5*, 125–155.

National Research Council. (2009). Strengthening Forensic Science in the United States: A Path Forward. National Academies Press.

National Commission on Forensic Science. (2015). *Ensuring That Forensic Analysis Is Based Upon Task-Relevant Information*. Department of Justice and Department of Commerce.

Nucci, Guilherme de Souza. (2024). *Código de processo penal comentado*. (23ª ed.). Forense.

Pacelli, Eugênio; Fischer, Douglas. (2016). *Comentários ao Código de Processo Penal e sua jurisprudência*. (8ª ed.). Atlas.

Páez, Andrés. (2021) Los sesgos cognitivos y la legitimidad racional de las decisiones judiciales. Em F. J. Arena; P. Luque; D. J. Moreno Cruz (Eds.). *Razonamiento Jurídico y Ciencias Cognitivas* (187-222). Universidad Externado de Colombia.

Prado, Geraldo. (2024). *Curso de processo penal: tomo I: fundamentos e sistema*. Marcial Pons.

President'S Council of Advisors on Science and Technology. (2016). *Report to the President. Forensic science in criminal courts: ensuring scientific validity of feature-comparison methods*. Executive Office of the President.

Reid, Melanie. A. (2018). CSI Story: the past, present, and future of crime scene collection and what litigators need to know. *Wake Forest Journal of Law & Policy, 8*(2), 409–454.

Salaverría, Juan Igartua. (2021). *Indícios, Duda Razonable, Prueba Científica: Perspectivas sobre la prueba en el proceso penal*. Tirant lo blanch.

Tonini, Paolo. (2004). Direito de defesa e prova científica: novas tendências do processo penal italiano. *Revista Brasileira de Ciências Criminais,12*(48), 194-214.

Tversky, Amos; Kahneman, Daniel. (1974). Judgment under Uncertainty: Heuristics and Biases. *Science*, *185*(4157), 1124-1131.

United States Department of Justice, Office of the Inspector General (2006). *A Review of the FBI's Handling of the Brandon Mayfield Case*.

Vázquez, Carmen. (2023). *Guía sobre el contenido de los informes periciales y su impacto en el debido proceso*. Consejo de la Judicatura Federal y Escuela Federal de Formación Judicial.

Vázquez, Carmen. (2022a). Las comunidades expertas y los sesgos cognitivos de los peritos. Em C. Vázquez, (Coord.). *Manual de Prueba Pericial* (43-94). Suprema Corte de Justicia de la Nación.

Vázquez, Carmen. (2022b). Los desacuerdos entre peritos y la junta pericial. Em C. Vázquez, (Coord.). *Manual de Prueba Pericial* (95-142). Suprema Corte de Justicia de la Nación.

Vázquez, Carmen. (2021). *Prova pericial: da prova científica à prova pericial*. JusPodivm.

Vieira, Renato Stanzola. (2023). *Controle da prova penal: obtenção e admissibilidade*. (2ª ed.). Revista dos Tribunais.

Leis e Jurisprudência

Decisões de Tribunais Brasileiros

Agravo Regimental no Habeas Corpus do Superior Tribunal de Justiça (Brasil) 898.724/SP (5ª Turma), de 12 de dezembro de 2024.

Apelação Criminal do Tribunal de Justiça do Estado de São Paulo (Brasil) 0251309- 33.2010.8.26.0000 (4ª Câmara de Direito Criminal), de 1º de junho de 2011.

Recurso em Habeas Corpus do Superior Tribunal de Justiça (Brasil) 200.979/SP (6ª Turma), de 15 de dezembro de 2025.

Leis e Diretrizes Normativas

Brasil. Lei nº 12.030, de 17 de setembro de 2009. Dispõe sobre as perícias oficiais e dá outras providências. (Diário Oficial da União, 18 de setembro de 2009). *https://www.planalto.gov.br/ccivil_03/_ato2007-2010/2009/lei/l12030.htm*

Código de Processo Civil (Brasil). (Diário Oficial da União, 17 de março de 2015). *https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2015/lei/l13105.htm*

Código de Processo Penal (Brasil). (Diário Oficial da União, 13 de outubro de 1941). *https://www.planalto.gov.br/ccivil_03/decreto-lei/del3689.htm*.

Decreto nº 42.847 do Estado de São Paulo, de 9 de fevereiro de 1998, dispõe sobre a estrutura organizacional da Superintendência da Polícia Técnico-Científica e dá providências correlatas. Diário Oficial do Estado de São Paulo, 10 de fevereiro de 1998. *https://www.al.sp.gov.br/norma/6257*

Quaestio facti. Revista Internacional sobre Razonamiento Probatorio
Quaestio facti. International Journal on Evidential Legal Reasoning
Sección: Ensayos
2026 I 11 pp. 45-62
Madrid, 2026
DOI: 10.33115/udg_bib/qf.i11.23242
Marcial Pons Ediciones Jurídicas y Sociales
© Renzo Cavani
ISSN: 2604-6202
Recibido: 23/01/2026 | Aceptado: 29/03/2026 | Publicado *online*: 29/05/2026
Editado bajo licencia Reconocimiento 4.0 Internacional de Creative Commons

«CARGA» DE LA PRUEBA: DEL RÉQUIEM A LA OBERTURA*

Renzo Cavani**
Pontificia Universidad Católica del Perú
renzo.cavani@pucp.edu.pe

RESUMEN: El presente artículo examina críticamente la noción de «carga de la prueba», cuestionando su formulación tradicional como carga subjetiva. Se argumenta que dicho concepto carece de precisión técnica, proponiéndose en su lugar la categoría de «condición contribuyente» para explicar la influencia de la actividad probatoria en el resultado del proceso. Asimismo, se defiende la autonomía conceptual de la «carga de aportación probatoria», identificándola como una auténtica carga procesal cuyo cumplimiento es determinante para la admisión formal de los medios de prueba y la evitación del rechazo liminar de la demanda. El trabajo finaliza presentando las conclusiones derivadas de este análisis.

PALABRAS CLAVE: carga procesal; carga de la prueba; carga de aportación probatoria; condición contribuyente; Georg Henrik Von Wright.

«BURDEN» OF PROOF: FROM REQUIEM TO OVERTURE

* Este texto se basa parcialmente en Cavani (2019) y, con algunas correcciones, en Cavani (2023). Una versión preliminar fue presentada en las III Jornadas Nacionales de Derecho Probatorio (Santiago de Chile, 9 y 10 de agosto). Agradezco a los profesores Jordi Ferrer Beltrán, Carmen Vázquez y Diego Dei Vecchi por sus comentarios críticos durante dicho evento. Mi gratitud se extiende también a los profesores Jordi Nieva Fenoll, Artur Carpes, Vélgika Rafaele, Fernando Medina y Ravi Peixoto por sus observaciones a una versión preliminar de este paper; así como a los/as revisores anónimos.

** Profesor ordinario asociado en la Pontificia Universidad Católica del Perú. Doctor (Universitat de Girona) y Magíster (Universidade Federal do Rio Grande do Sul). Miembro ordinario de PRODEJUS-PUCP. CEO y socio cofundador de Evidence Lab.

ABSTRACT: This article critically examines the concept of «burden of proof», questioning its traditional formulation as a subjective burden. It is argued that said concept lacks technical precision; instead, the category of «contributory condition» is proposed to explain the influence of evidentiary activity on the outcome of the proceedings. Furthermore, the conceptual autonomy of the «burden of evidentiary production» is defended, identifying it as a genuine procedural burden, the fulfillment of which is decisive for the formal admission of evidence and the avoidance of a dismissal *in limine* of the complaint. The paper concludes by presenting the findings derived from this analysis.

KEYWORDS: procedural burden; burden of proof; burden of evidentiary production; contributory condition; Georg Henrik Von Wright.

SUMARIO: 1. INTRODUCCIÓN.— 2. CARGA, CARGAS PROCESALES, CARGA PROBATORIA.— 3. «CARGA» DE LA PRUEBA: VIEJAS Y NUEVAS CRÍTICAS.— 4. LA APORTACIÓN DE PRUEBA COMO CONDICIÓN CONTRIBUYENTE.— 5. CARGA DE APORTACIÓN PROBATORIA: UNA DEFENSA DE SU AUTONOMÍA CONCEPTUAL.— 6. CONCLUSIONES.— REFERENCIAS BIBLIOGRÁFICAS.

1 INTRODUCCIÓN

El concepto de «carga de la prueba» ha sido históricamente problemático y, a menudo, aceptado dogmáticamente sin un análisis conceptual riguroso que lo vincule adecuadamente con la teoría general de las situaciones jurídicas (específicamente, el concepto de «carga»).

Así, recogiendo la exposición tradicional de este concepto, algunas críticas esbozadas en los últimos años y, además, algunas réplicas, el presente trabajo tiene por objeto contribuir con el debate agregando (o, acaso, afinando) algunos argumentos críticos contra la noción tradicional de «carga» de prueba, demostrando sus falencias para encuadrar dentro de la categoría de «carga».

A lo largo de estas líneas, se insistirá en que la llamada «carga de la prueba subjetiva» no constituye una verdadera regla técnica, dado que la aportación de prueba no garantiza por sí sola el resultado de vencer en la causa; por ello, se recurre a Von Wright a fin de reconducir la noción de estímulo o incentivo de aportar pruebas (algo importante para la doctrina tradicional) al concepto de condición por la de *condición contribuyente*; esto es, una condición necesaria de una de las condiciones suficientes posibles para vencer en la causa.

Sin embargo, lejos de desechar toda utilidad del concepto de «carga» en esta discusión, se defenderá la autonomía conceptual de la «*carga de aportación probatoria*». Se sostendrá que esta figura sí cumple con los requisitos de una carga genuina, pues su resultado normativo específico ya no sería la victoria en el litigio, sino, entre otros, la admisión formal de los medios probatorios y la evitación del rechazo liminar de la demanda.

Se finalizará con algunas conclusiones.

2. CARGA, CARGAS PROCESALES, CARGA PROBATORIA

El concepto de «carga de la prueba» no deja de ser realmente problemático. Esto se ha venido evidenciando desde los tratados clásicos[1] hasta las contribuciones doctrinarias más modernas: algunos proponiendo la eliminación completa de su uso[2]; otros, justificándolo[3].

Como es común, desde una perspectiva conceptual, para hablar de «carga de la prueba» es bastante útil buscar conceptos más generales, como es el caso de *cargas procesales*. Inclusive, se podría partir desde un concepto aún más general: la *carga.* Hay ríos de tinta escritos sobre el tema[4], pero es posible partir del entendimiento de que una *carga* presupone un *ejercicio de libertad* del que goza alguien a fin de obtener un *resultado favorable*, sujeto a la realización de un cierto comportamiento[5].

Vincular la carga con «imperativo de conducta» (como hacía Goldschmidt) o con «peso»[6] oscurece el aspecto central del concepto, a saber: la *libertad.* El titular de una carga la ejercita si así lo quiere, con el fin de obtener una ventaja o un resultado positivo[7]. Un buen ejemplo sería el caso de un contribuyente titular de un bien inmueble que debe pagar impuestos a la municipalidad en que se sitúa su bien (en el Perú se denominan «arbitrios»). Así, si el contribuyente es moroso, la municipalidad anuncia que se puede acoger a un descuento del 50 por cien si paga antes de una cierta fecha. Aquí tenemos la carga: el moroso tiene la *libertad* de acogerse o no a dicho descuento. Si quiere, paga menos dinero; si quiere, paga más[8].

---

[1] Augenti (1932); Rosenberg (2002); Micheli (1966); Verde (1974).

[2] Paula Ramos (2015: pp. 53 ss.; 2015a); Nieva Fenoll (2019, 2024, 2025); Ferrer Beltrán (2019).

[3] Mitidiero (2020); Vogt Geisse (2025).

[4] Cfr. Gavazzi (1970). La posición de Carnelutti sobre el tema se extiende a lo largo de muchas de sus obras, ver Augenti (2000: pp. 217-222).

[5] Quizá una definición más precisa sea la siguiente: una *carga* es una situación jurídica subjetiva cuya realización depende de su titular para la obtención de un determinado beneficio, sin que medie ningún tipo de comportamiento correlativo por la parte pasiva y sin coerción por el ordenamiento jurídico. No profundizaré en este tema por ahora. Para una problematización, cfr. Gavazzi (1970: pp. 13 ss.), Paula Ramos (2020: pp. 74-76).

[6] Por ejemplo, Falcón conceptualiza de la siguiente manera la *carga de la prueba:* «La carga de la prueba es el imperativo, o el peso que tienen las partes de recolectar las fuentes de prueba y activarlas adecuadamente para que demuestren los hechos que les corresponda probar a través de los medios probatorios y sirve al juez en los procesos dispositivos como elemento que forma su convicción ante la prueba insuficiente, incierta o faltante» (2003: p. 247).

[7] Para un estudio sobre el tema, véase Fernández López (2006); Paula Ramos (2020: pp. 60 ss.); Rodríguez Álvarez (2020: pp. 127 ss.).

[8] Como bien indica Rodríguez Álvarez (2020), «que la carga no es que se cumpla o se incumpla, sino que se observa o no se observa» (p. 141) y, hablando específicamente sobre la carga procesal, «el hecho de que no haya sanción no implica que no exista una consecuencia jurídica frente a su inobservancia (o, en ocasiones, su defectuosa observancia), pues evidentemente la hay. Dicha consecuencia constituye una modificación de una determinada situación jurídica: el empeoramiento de la posición procesal del

Pero la carga no se agota aquí. Si así fuera, no habría distinción con otras situaciones jurídicas como la facultad o la propia libertad. El otro aspecto central de la carga (y esto es relevante para la «carga de la prueba») es que la satisfacción producto de la ejecución del comportamiento vinculado a la situación jurídica *no depende de ningún otro comportamiento*. Volvamos a nuestro ejemplo: si el contribuyente moroso titular de la carga decide acogerse al descuento, entonces basta con pagar antes de la fecha indicada y, con ello, se acoge al descuento del 50 por cien.

Si esto es carga, entonces, ¿qué sería una *carga procesal*? Sería no otra cosa que una carga que se ejerce en el marco de un proceso o que, al menos, tiene efectos respecto del proceso[9]. Por supuesto, existen muchas *cargas procesales* (en plural). Se puede mencionar, por ejemplo, la *carga de individualizar el pedido o petitorio*, la *carga de comparecer el proceso*, la *carga de recurrir*, la *carga de identificar los bienes para una futura ejecución*, etc. Todas estas acciones son ejercicios de libertad porque presuponen un espacio de autonomía frente al Estado e, inclusive, frente a la contraparte, respecto del cual no puede haber interferencias. Asimismo, el resultado producto del comportamiento vinculado al ejercicio de la carga es variado: delimitar el objeto del proceso y la cuestión controvertida, en el caso de la carga de individualización del pedido; no ser declarado rebelde, ser notificado en el domicilio indicado y establecer patrocinio legal, en el caso de la carga de comparecer al proceso; impedir que la resolución sea firme, en el caso de la carga de recurrir; evitar que pueda haber embargos respecto de bienes valiosos para el deudor, en el caso de la carga de identificar bienes, etc.

Dentro de las cargas procesales se encuentran las *cargas probatorias*. Estas serían, en un sentido muy genérico, aquellas que otorgan la *libertad* de diseñar las actuaciones o la estrategia procesal en cuanto a los hechos alegados y al caudal probatorio que se destina a acreditarlos. Existen varias cargas probatorias, tales como la *carga de alegar los hechos* (o, también, la incorporación de los elementos fácticos en la causa de pedir[10]) y la *carga de argumentar a partir del material fáctico incorporado por las partes*. La ventaja que se obtiene, sin mediar ningún otro comportamiento, es precisamente estructurar la causa de pedir y, con ello, el *thema probandum* y, además, incluir información para el debate probatorio y la promoción del derecho al contradictorio. Todas estas son, pues, auténticas cargas, aun cuando no necesariamente sean vistas de esta manera.

La «carga» de la prueba sería una especie de carga probatoria; aunque más adelante veremos que hay diversos problemas conceptuales para encajarla siquiera como *carga*.

---

sujeto, en la medida en que se reducen las expectativas de obtener un resultado favorable a sus intereses» (pp. 141-142).

[9] Ello remonta al tema de saber cuándo un acto puede calificarse como *procesal*, dependiendo de diversos criterios, como la sede, los efectos o los sujetos. cfr. Conso (1969: p. 187); Calmon de Passos (2009: pp. 43 ss.); Didier & Nogueira (2012: pp. 33-35).

[10] Destaca esta carga Rivera Morales (2010: p. 130), aunque reconoce que la «carga» de la prueba sería, más bien, una facultad.

3. «CARGA» DE LA PRUEBA: VIEJAS Y NUEVAS CRÍTICAS

Desde una explicación bastante tradicional[11], se dice que la así llamada «carga de la prueba» consiste en una *regla de juicio* que funciona a modo de desempate, a fin de determinar al vencedor en los casos en que existe insuficiente probanza de uno o más hechos esenciales (los así denominados hechos constitutivos, extintivos, modificativos e impeditivos)[12]. Pierde, pues, quien tiene la titularidad de la carga de probar el hecho que, al final, no se llegó a probar; esto es, si el juez determinó que la hipótesis probatoria no quedó *suficientemente* acreditada. Si el demandante alega que se celebró un contrato verbal, pero no lo llega a probar, entonces le correspondería perder[13]. Ya en el caso del demandado, si es que alega que existió pago y no lo llega a probar, entonces vencería el demandante (porque el art. 1229 del Código Civil peruano indica que la prueba del pago corresponde a «quien pretende haberlo efectuado»[14]), pero siempre que este pruebe la existencia de la deuda.

En materia laboral, según la legislación peruana, si el demandante-trabajador alega el pago, el incumplimiento de normas, el incumplimiento del contrato, la vigencia del contrato o su exigibilidad, el demandado-empleador debe probar que hubo pago, que se cumplieron las normas, que el contrato se extinguió o que ya no es más exigible (art. 23.4 de la Ley Procesal del Trabajo[15]). Ya en materia penal, al Ministerio Público le corresponde demostrar que está probada la hipótesis inculpatoria (art. IV.1 del Código Procesal Penal peruano[16]); si esto no llega a estar *suficientemente* corroborado, entonces, aplicando la regla de juicio denominada presunción de inocencia[17], el acusado queda absuelto (art. II.1)[18].

---

[11] Para una aproximación histórica y comparatística, cfr. Micheli (1966: pp. 3-58); Comoglio (2013); y, recientemente, Nieva Fenoll (2025). Para una evolución de la institución en el marco de la doctrina alemana, cfr. Prütting (2010).

[12] Para una exposición dogmática sobre la carga de la prueba, cfr. Didier Jr. (2017).

[13] En esta hipótesis se aplicaría la carga de la prueba, por supuesto, siempre que el demandado no pruebe que pagó. Como veremos más adelante, hay diversas razones por las que, a pesar de que se cumpla con la carga, al final no se produciría el resultado esperado. Precisamente esto hace que sea inadecuado emplear el concepto de «carga» para describir este fenómeno.

[14] Decreto Legislativo N° 295, de 25 de julio de 1984, Código Civil.

[15] Ley N° 29497, de 15 de enero de 2010, Ley Procesal del Trabajo.

[16] Decreto Legislativo N° 957, de 29 de julio de 2004, Código Procesal Penal.

[17] La «presunción de inocencia», en rigor, es una *presunción aparente* o «verdad interina». Esto quiere decir que se admite prueba en contrario y la «carga de la prueba» se sitúa en aquel que no es favorecido por la presunción, pero, a diferencia de una presunción en sentido estricto, no existe hecho base. De cualquier manera, se trata de una norma que contiene una obligación para el juez de fallar a favor del acusado si no se supera el estándar de culpabilidad.

[18] De estos ejemplos vemos que la titularidad es asignada por el legislador (distribución estática), que puede corresponder a la parte que alega el hecho (distribución estática fija) o la contraparte que no ha alegado (distribución estática invertida o «*inversión* de la carga de la prueba»). Excepcionalmente, el propio legislador puede otorgar al juez el poder de ser él quien la distribuya (distribución *dinámica*), como ocurre en el caso del 33 de la Ley del Proceso Contencioso Administrativo peruana, específicamente

Siempre según la explicación tradicional, esta regla de juicio, o regla de desempate, es la que incentiva o estimula a las partes a que puedan aportar pruebas a fin de que no se produzca la consecuencia desfavorable de perder la causa[19]; o sea, que no se vean perjudicadas o que, al menos, no haya riesgo de perjuicio[20]. Se habla aquí de una *regla de instrucción* que informaría a las partes sobre cómo estructurar u optimizar su actividad probatoria[21], o hacer al menos que se «esfuerce» para evitar ese resultado desfavorable[22]. Empero, dado que una situación jurídica como la carga requiere que esté reconocida en una norma, estaríamos frente a la siguiente situación: a partir de las *mismas* disposiciones normativas que expresan las normas de obligación para un juez sobre cómo fallar en una controversia si hay falta de probanza, también podrían interpretar *normas* que consagran genuinas *cargas*. Pero estas cargas no sólo podrían limitarse a «estimular», sino a regular algo muy preciso: *si* se realiza una acción *X* (aportar prueba, estructurar una cierta estrategia, etc.), *entonces*, sin mediar otro comportamiento, se alcanzará el resultado *B* (ganar la causa, disminuir el riesgo de perderla, etc.). Sólo así se podría mantener en pie dicho concepto[23].

Contra esta forma de entender las cosas se pueden tejer algunas críticas.

---

en los casos en que, «por razón de su función o especialidad la entidad administrativa está en mejores condiciones de acreditar los hechos». Así también ocurre en muchas otras legislaciones procesales. Nótese bien: rigurosamente, la distribución dinámica (o «dinamización de la carga de la prueba») no se da cuando el juez, *sin* autorización expresa, decide alterar la titularidad de las cargas de corroboración. Esto no es más que una muy discutible derrotabilidad de la norma que dispone la distribución estática a través del «control de constitucionalidad» u otra técnica interpretativa. Esto es lo que defiende, por ejemplo, Peyrano (2004) y, con mayor sofisticación, Carpes (2010) y Buril de Macêdo & Peixoto (2014). Ambas posiciones han recibido una dura crítica por Paula Ramos (2015: pp. 47 ss., 75 ss.) con otros fructíferos debates posteriores. En el caso del CPC brasileño de 2015 se llegó a consagrar la distribución dinámica, art. 373 § 1 (cfr. Carpes, 2015 y 2017); y una opción similar ha optado el Proyecto de Reforma del CPC peruano, publicado en 2021. En mi caso, tengo muchísimos reparos contra esta distribución realizada por el juez «para el caso concreto» pues, en la misma línea de Costa (2017: pp. 137 ss.), la decisión contraria al demandante/acusador por falta de probanza sí que tiene sustento constitucional. Para críticas contra la carga dinámica de la prueba, cfr. Falcón (2003: pp. 278 ss.); Ferrer Beltrán (2019), Paula Ramos (2020: pp. 95 ss.).

[19] Es conocida la lección de Rosenberg (2002: p. 37), quien, apartándose de una fuerte tradición doctrinaria enfocada en la «dimensión subjetiva», identifica que el problema de la «carga de la prueba» se presenta cuando hay una situación de falta de probanza. No obstante, opina que no se debiera eliminar el *concepto* de carga subjetiva (sin interesar el nombre que se le dé) pues, en el marco de un proceso dominado por el principio dispositivo, las alegaciones de hecho y la producción de la prueba no es sólo una «consecuencia práctica para prevenir el peligro de la pérdida del proceso, o algo que sólo se aconseja a las partes, sino que una consecuencia jurídica» (p. 38). Más bien, se corresponde con el deber del juez de requerir a la parte que aporte medios de prueba. Esta construcción teórica, enfatizando en el papel del principio dispositivo e inquisitivo, fue puntualmente analizada por Micheli (1966).

[20] Paula Ramos (2020: p. 56).

[21] Carpes (2017).

[22] Rosenberg (2002: pp. 31 ss.).

[23] Más adelante veremos que se buscó proponer el concepto de «carga imperfecta» para dar cuenta del fenómeno en que, aun cumpliendo la carga, no se daría el resultado prometido.

En primer lugar, decir que la «carga de probar en sentido *objetivo*» *es* una regla de juicio dirigida al juez significa incurrir en una seria contradicción. Si *es* carga, entonces es titularizada por una parte; si, por el contrario, *es* una regla dirigida al juez se crea una situación jurídica titularizada por este —que, en rigor, sería un *deber*— y no por la parte. Desde una perspectiva lógica, de un deber no se sigue una carga[24]; y tampoco se podría decir que el deber, como posición jurídica fundamental que es, puede tener a la carga como un opuesto o como un correlativo[25]. O sea, conceptualmente hablando, la existencia del deber del juez de resolver en un cierto sentido en casos de falta de probanza, no implica la existencia de una carga de la parte. En simple: de las disposiciones normativas que regulan las reglas de desempate *necesariamente* hay un *deber*, pero no *necesariamente* alguna carga.

Si asumimos esta premisa, la así llamada «carga de la prueba objetiva» (esto es, la regla de juicio) no tiene ninguna *relación conceptual* con la así llamada «carga de la prueba subjetiva». La pregunta «qué parte *debiera* aportar prueba al procedimiento» a partir de los incentivos que se generarían en ellas a fin de no perder la causa[26] no se responde con la pregunta «¿quién pierde si no hay prueba suficiente?». Esta última pregunta se responde de la siguiente manera: a partir del deber del juez de resolver de una determinada manera si, luego de valorar, constata la falta de suficiencia de una cierta hipótesis probatoria. Por ello, ya desautoriza poder hablar de «dimensiones objetiva y subjetiva» de *un mismo concepto*, como proponía Rosenberg[27] y como es defendido por buena parte de la doctrina. El *aggiornamiento* terminológico y conceptual se impone[28].

En segundo lugar, quedaría saber si existe genuinamente alguna *carga* en la así llamada *dimensión subjetiva* de la carga de la prueba. Si buscamos entenderla como una *regla técnica* dirigida a las partes a fin de que aporten medios de prueba (pues, de no hacerlo, perderían la causa), me parece que estamos frente a un error conceptual. En efecto, una regla técnica no es una prescripción, pues depende de la *voluntad* de la parte de realizar un cierto comportamiento para obtener un cierto resultado como

---

[24] Ferrer Beltrán (2019: p. 60).

[25] En efecto, según Hohfeld (2004), el opuesto jurídico de un *deber* sería un *privilegio*, en el sentido de que, si X tiene el derecho de entrar en un inmueble, *no tiene* el deber de no quedarse fuera, como sí lo tendrían otras personas (p. 53). Por su parte, el correlativo jurídico del *deber* sería el *derecho subjetivo* (pp. 49 ss.), en el sentido de que para satisfacer un derecho se requiere del cumplimiento de un deber.

[26] Ferrer Beltrán (2019: pp. 59 ss., 69-74).

[27] Cfr. Prütting (2010).

[28] Nieva Fenoll (2024 y 2025), aunque el autor propone una eliminación completa de la *institución* (para una crítica a dicho autor, véase Vogt Geisse, 2025). Más convincente es Paula Ramos (2021: p. 252) cuando afirma que siempre se necesita una regla de desempate (no interesa cómo se le llame), y ello no se logra simplemente con encomendar al juez a que realice la valoración de la prueba: siempre se necesita una *regla* que establezca una acción respecto de qué hacer en caso de desempate. Por su parte, como bien señala Rodríguez Álvarez (2020), «no nos encontramos ante la teoría de la carga de la prueba, sino ante la teoría de la regla de juicio. Sólo razones de inercia histórica podrían explicar que se mantenga esta nomenclatura» (p. 153).

condición necesaria, como sería el caso, por ejemplo, de la siguiente proposición: «Si desea encender la televisión, apriete el botón rojo»[29]. No obstante, como acabamos de ver, la doctrina tradicional entiende que, en el caso de la carga de la prueba en sentido subjetivo, el comportamiento «aportar prueba» tendría que llevar al resultado «vencer en la causa» o, inclusive, una decisión probatoria favorable[30] (de lo contrario, no podría llamársele una *carga*). Empero, el comportamiento de la parte titular de la situación jurídica no conduce por sí sólo a la situación ventajosa, como vimos en el caso del contribuyente moroso. En este último caso, su sola acción le conduce a obtener el beneficio vinculado al ejercicio de su carga. Ya en el caso de la «carga de la prueba», para que una parte venza en la causa dependerá de una serie de cuestiones, entre ellas que el juez considere que sus pruebas aportan información fiable, en calidad y cantidad, y que la hipótesis probatoria efectivamente ha superado un cierto estándar de prueba[31].

En tercer lugar, como ha sido destacado, tenemos también el «principio de adquisición de prueba», norma que suele pertenecer a nuestros sistemas jurídicos, que regula que las pruebas ofrecidas por las partes pasan a formar parte del proceso y el debate probatorio se desarrollaría a partir de ellas, sin perjuicio de quien las aportó. Por ello, el juez valora *todas* las pruebas admitidas y construye inferencias probatorias, tanto individual como conjuntamente, sin interesar de qué parte provino la prueba. Esto quiere decir no otra cosa que lo siguiente: una parte puede aportar un medio de prueba respecto del cual, a la postre, pueda llevar a la conclusión que el hecho que ella pretendía acreditar, en realidad, no se produjo o, también, que el derecho alegado no le asiste[32].

Este argumento se dirige a controvertir la existencia *conceptual* de la situación jurídica «carga» de la prueba — específicamente la subjetiva—; pero, nótese bien, no implica que no pueda existir un *efecto pragmático* de las reglas dirigidas al juez sobre cómo fallar ante la falta de prueba. En efecto, no se puede negar que la distribución de la «carga» de la prueba respecto del comportamiento de las partes respecto a cómo

---

[29] Todo según Ferrer Beltrán (2019: pp. 61-62), con apoyo en Gavazzi y, especialmente, en Von Wright (1963: pp. 9 ss.), respecto del concepto de *regla técnica* y *proposición anankástica*.

[30] No obstante, como dice Nieva Fenoll (2025: p. 438): «Lo que sucede en los sistemas de *Civil Law* es que el juez falla en contra de un litigante, no porque no haya aportado prueba, sino porque la —poca— prueba que ha aportado no le convence, que es muy diferente, lo que se traduce en una pérdida del proceso como consecuencia de que la valoración de la prueba le es adversa al litigante, y no porque "no haya cumplido la carga de la prueba", al contrario de lo que se dice tantísimas veces en las sentencias, precisamente para que el juez se ahorre la motivación de la valoración de la prueba, que es más compleja».

[31] Mitidiero (2020) defiende que la carga sería una situación jurídica una vez que es *libremente ejercitada*, sería «*capaz probablemente de promover* el resultado pretendido por el ordenamiento jurídico y por el titular de la carga» (p. 25), con lo que se justificaría la distinción entre «carga perfecta» (cuando el ejercicio lleva al resultado) y «carga imperfecta» (cuando el ejercicio lleva *probablemente* al resultado). Sobre esto volveré en el acápite 5.

[32] Conforme: Paula Ramos (2020: p. 96); Ferrer Beltrán (2019: pp. 61 ss., 71, nota 40).

desempeñar su actividad probatoria en el proceso puede, de alguna manera, orientar su comportamiento[33].

De hecho, este es un efecto bastante normal en las normas jurídicas[34]: por ejemplo, el deber de no causar daño puede llevar a que, en una ciudad como Lima o Río de Janeiro, uno deba manejar con sumo cuidado, tomando acciones de mayor prevención. Ahora pensemos en este ejemplo: la norma del proceso laboral peruano indica que al trabajador-demandante sólo le corresponde alegar y probar la existencia del vínculo laboral, mientras que al empleador-demandado le corresponde probar que dicho vínculo responde a un contrato temporal y no a uno de plazo indeterminado. Según la legislación laboral, el contrato temporal debe ser por escrito (este es un deber, sujeto a sanción). Pues bien, esta regla de juicio puede perfectamente generar —y, de hecho, podríamos decir que es lo que *normalmente* ocurre— un *incentivo* en el empleador de contar no sólo con los contratos firmados, sino debidamente ordenados y sistematizados. Y también esto tendría un *incentivo* para saber, desde una perspectiva estratégica, en qué casos colocar más recursos en honorarios de abogados, según la contingencia laboral. Pues bien, mi punto es que estos incentivos, si bien sí existen, no responden necesariamente al concepto de *carga*, ni tampoco existe necesidad conceptual de encuadrar tales comportamientos en dicho concepto[35].

Nótese, además, cómo este mismo ejemplo podría generar otro tipo de incentivos. Imaginemos que el trabajador-demandante, consciente de la existencia de la regla de juicio y, por tanto, de lo que le corresponde alegar y probar, decide ir más allá: contrata a un buen abogado y busca demostrar que su vínculo es, en efecto, de plazo indeterminado, aportando una grabación del empleador respecto de las negociaciones previas al cierre del contrato. ¿Podría hablarse aquí de alguna «carga» o, si la hubiere, de que habría sido cumplida?

En cuarto lugar, y conectando con el argumento anterior, si acaso estuviésemos frente a una «carga» respecto de llevar adelante una cierta estrategia probatoria me-

---

[33] Esto algo que reclama Mitidiero para defender el concepto de «carga» de la prueba (2020: p. 28). Por razones expuestas en el texto, sin embargo, pienso que se puede no sacrificar ese «aspecto subjetivo» pero sin tener que comprometerse con el concepto «carga»; salvo que, por supuesto, se convenga seguir haciendo homenaje a la tradición.

[34] Conforme: Ferrer Beltrán (2019: p. 70).

[35] Me parece que este también es el argumento de Paula Ramos (2021: p. 238, nota 100). Nótese, además, que pueden existir acicates o empujones (*nudges)* plasmados en ciertas normas jurídicas para promover ciertos comportamientos en las personas. No obstante, los *nudges*, si bien presuponen un ejercicio de autonomía (seguir o no una conducta, aunque *empujando* a que se adopte alguna), no necesariamente están consagrados en normas que regulan cargas; también puede hablarse de derechos, facultades, privilegios, etc. En suma, que deba descartarse el término «dimensión subjetiva de la carga de la prueba» no quiere decir que no existan *incentivos prácticos.* Vogt Geisse (2025) enfatiza también en la existencia de estos incentivos (que resultan del ejercicio de la autonomía de la voluntad), pero esto le lleva a concluir que debiera mantenerse el *concepto* de «carga de la prueba subjetiva» (pp. 225 ss.). Tal como veremos en el texto, en efecto, es necesario dar cuenta de estos incentivos (fenómeno que sí existe), pero el concepto de «carga» no sirve para ello.

diante la aportación de pruebas, no se indica exactamente qué tipo de pruebas se tendrían que presentar, cómo se tendría que alegar respecto de ellas ni cómo desarrollar esta estrategia. De nuevo, esto contrasta con el ejemplo del contribuyente moroso: «Pague usted antes de tal fecha, y se acogerá a un descuento, por lo que pagará menos». En el caso de la «carga» de la prueba solamente existe un conjunto de reglas que nos dicen que si se alega un hecho correspondería ejercitar actividad probatoria tendiente a probar dicha alegación (o, en caso de inversión, si la contraparte alega un hecho correspondería contraprobar dicha afirmación). Vistas así las cosas, se trata de una *regla increíblemente genérica* y ambigua como para guiar con precisión la conducta de las partes; en efecto, si el comportamiento no está suficientemente especificado no parecería suficiente para cumplir con una característica básica del concepto de carga.

Frente a estos inconvenientes conceptuales que ofrece el concepto «carga» de la prueba, la procesalística ha propuesto entenderla como una *carga imperfecta*[36]. El razonamiento es particularmente simple: si la *carga perfecta* lleva necesariamente a la producción del resultado, la carga *imperfecta* o *incompleta* no llevaría necesariamente a este[37], sino que *tendería* a producirlo. Por ejemplo, Micheli (con base en Golschmidt y Allorio), aunque siendo crítico y escéptico, recoge las opiniones de la doctrina, y entiende a la carga como una correlación entre el ejercicio de una «facultad» y la consecuencia de un resultado; pero este puede tener algunas atenuaciones en la medida que: *a)* el resultado no depende de su propia actividad probatoria (*carga incompleta*)[38] o *b)* otro sujeto (el juez) «tendría el poder de integrar el ejercicio insuficiente de la facultad»[39], en este caso mediante poderes instructorios (*carga imperfecta*)[40].

---

[36] Críticamente, cfr. Micheli (1966). Defienden este concepto Carpes (2017); Mitidiero (2020: pp. 24-25). Para una exposición crítica, cfr. Paula Ramos (2020: pp. 77 ss.). Mostrando su origen en Goldschmidt, ver Rodríguez Álvarez (2020: pp. 143 ss.).

[37] A diferencia de lo que entiende Mitidiero (2020: p. 24), Gavazzi (1970) no estaría asumiendo la posibilidad de que exista una «carga imperfecta». Más bien lo que dice, al analizar la teoría de la invalidez de Conso (pp. 136-137), es que los actos imperfectos (esto es, viciados) sujetos a la regla de la conservación pueden producir efectos equiparándose a un acto válido (p. 138). Así, la relación que Gavazzi ve entre carga e invalidez es doble: 1) de las nociones de validez/invalidez se deduce el concepto de carga (p. 138); y 2) del régimen de equiparación entre acto válido e inválido en cuanto a los efectos, se puede realizar una interpretación sistemática respecto del funcionamiento de la carga que sintetizaría de la siguiente manera: «Si usted quiere relevancia para el acto jurídico *X*, celébrelo de la forma *Y*; pero si lo celebra de la forma *W* podría ser inválido, aunque si hay subsanación de *X*, podrá ser relevante como si lo hubiera celebrado de la forma *Y*». Una explicación de esta tesis de Conso puede encontrarse en Cavani (2014: pp. 327 ss.).

[38] Micheli (1966: pp. 148-149).

[39] Micheli (1966: pp. 61-62, nota 6).

[40] Micheli es muy crítico y escéptico con ambas categorías. En primer lugar, según el autor, la *carga imperfecta* reúne dos fenómenos diferentes: incertidumbre sobre la probanza de los hechos y la relación entre juez y partes; en segundo lugar, tanto la *carga imperfecta* como la *carga incompleta* confunde la fase de valoración de la prueba con la fase de decisión. Para Micheli, la carga de la prueba se restringe fundamentalmente a la *regla de juicio* evitando el *non liquet*; por lo que hablar de «carga imperfecta» o

No obstante, me parece que este no es el mejor camino, en gran medida porque si no se prevé un resultado específico para el cumplimiento de una carga, se pierde su propia normatividad, en el sentido de que no queda establecida la acción específica a adoptar. No sería, por tanto, una carga en sentido jurídico, sino nada más que una *una necesidad que se verifica en la práctica jurídica*[41].

4. LA APORTACIÓN DE PRUEBA COMO CONDICIÓN CONTRIBUYENTE

Una mejor explicación que recurrir al concepto de «carga» (y, aún más, al equívoco concepto de «carga imperfecta» o «carga incompleta») para dar cuenta del fenómeno del estímulo o del incentivo para realizar las acciones «aportar prueba», «tejer una estrategia probatoria» o «definir el camino probatorio a recorrer»[42] es hablar de una *condición contribuyente* para lograr el resultado esperado (que podría ser vencer en la causa o, inclusive, corroborar la hipótesis probatoria).

Me explico. Según Von Wright, una propiedad *A* es una *condición suficiente* de una propiedad *B* en tanto que siempre que *A* esté presente, *B* también lo está ($A \subset B$). Por su parte, una propiedad *A* es *condición necesaria* de una propiedad *B* en tanto que sin la presencia de *A* no estará presente *B* (pero se requiere de la presencia de otras condiciones, además de A, para que *B* sí lo esté); o sea, $B \subset A$. Finalmente, una propiedad *A* es una *condición contribuyente* de una propiedad *B* siempre que *A* sea *condición necesaria* de al menos una *condición suficiente* de *B*[43].

Así, tenemos las propiedades «aportar prueba» (*A)* y «vencer en la causa» *(B).* Como fue indicado, el aportar prueba (más allá que esto no resulte del todo claro, como ya se vio) no presupone que se venza en la causa, dado que la prueba puede ser irrelevante, ilícita o que se haya ofrecido información muy poco fiable (imaginemos un documento audiovisual mal conservado o con poca calidad técnica). Por su parte, vencer en la causa no presupone necesariamente haber aportado prueba, porque podría haber confesión de la contraparte, alguna presunción en favor de quien le correspondía aportar prueba o simplemente porque se valoró desfavorablemente una prueba ofrecida por una parte (producto del principio de adquisición de la prueba, como ya se vio). Así pues, «aportar prueba» no es condición necesaria ni tampoco condición suficiente de vencer en la causa.

¿Por qué digo, en todo caso, que podría ser una condición contribuyente? Dado que vencer en una causa *(B)* depende de una compleja conjunción de diversos factores, tenemos que *B* exige una *condición suficiente compleja,* porque *a)* son diversas

---

«incompleta» no contribuye en nada para esclarecer conceptualmente la función de decidir ante la falta de probanza (1966: p. 61, nota 6, pp. 173, 252).

[41] Esta crítica es de Paula Ramos (2020: pp. 78-80). Ver también Paula Ramos (2021: p. 259).

[42] Carpes (s/f), quien introduce la idea de «camino» (en portugués: «*roteiro*»).

[43] Von Wright (2001: pp. 66 ss., esp. 73 ss.).

condiciones suficientes y *b)* cada condición suficiente puede contar, a su vez, con diversas condiciones necesarias, esto es, para que *B* se produzca es *necesario* que estén presentes *B', B"… B*^{n}.

Imaginemos la siguiente condición suficiente S_1 = (hechos relevantes + norma aplicable + aportación de prueba + corroboración de hipótesis probatoria + suficiencia de hipótesis según un cierto estándar de prueba). Si se cumple, se vence en la causa ($S_1 \subset B$). Así, sin aportar prueba, no se puede lograr S_1. Aportar prueba es una condición necesaria de una de las tantas formas suficientes de lograr vencer en la causa. Y aquí reside un punto importante: aportar prueba *(A)* no sería una *condición contribuyente indispensable* precisamente porque hay otros caminos para poder ganar sin aportar prueba (S_2, S_3… *Sn,* en donde, por ejemplo, aportar prueba, no es una propiedad presente para conformar la condición suficiente S_2)[44]. Aquí, aportar prueba, en la teoría de Von Wright, sería una *condición contribuyente* (a secas) de una condición suficiente.

5. CARGA DE APORTACIÓN PROBATORIA: UNA DEFENSA DE SU AUTONOMÍA CONCEPTUAL

En el acápite 3, hemos visto que no se puede confundir la regla de juicio (deber titularizado por el juez) con alguna «carga» o alguna «dimensión subjetiva» de alguna carga. No obstante, esto no quiere decir que se deba descartar de plano el concepto de carga para dar cuenta de un fenómeno similar a lo que la doctrina tradicional viene aludiendo. Se trata de la *carga de aportación probatoria*.

A primera vista, podría pensarse que esta «carga de la prueba subjetiva» podría asociarse a la *carga de aportación*, pero creo que ello no es así. En efecto, podemos partir del esquema de la regla técnica explicada por Ferrer Beltrán, consistente en el comportamiento «aportar prueba», pero —aquí lo importante— ya no más con el resultado «vencer en la causa». Se requeriría contar con un resultado que se produzca sin la concurrencia de ningún otro comportamiento salvo el del titular de la carga. Se podría pensar, en primer lugar, en un resultado bastante simple: que dicha prueba ingrese *formalmente* al proceso (la fórmula en el lenguaje judicial podría ser la siguiente: «*téngase por ofrecidos los medios probatorios*», lo cual es algo que el juez indica *certificando* que las pruebas ingresaron conjuntamente con un acto postulatorio dentro del plazo indicado en ley). Aquí tenemos que la aportación probatoria es *condición necesaria* para que tales medios probatorios, por haber ingresado, sean oportunamente calificados para decidir su admisión o rechazo.

Pero no sólo ello. La carga de aportación podría tener una sólida justificación *dogmática*, vinculada en este caso a las particularidades del derecho positivo. Aquí

[44] Lo mismo podría decirse si hablamos del resultado «corroboración de hipótesis probatoria». Esto se verá al final del próximo acápite.

podemos advertir otro tipo de resultados generados por el cumplimiento de la carga. En efecto, si no se aportan medios de prueba que, para un juez, en el marco de un examen preliminar, le parezcan suficientemente relevantes para continuar con el procedimiento, entonces el proceso podría concluir inmediatamente (lo cual ocurre en la *preliminary hearing* del proceso penal estadounidense, *ex* Regla 5.1.f de las *Federal Rules of Criminal Procedure*[45]). De esta manera, en estos sistemas, la aportación de la prueba por el demandante o demandado sería una condición necesaria para no ver la demanda preliminarmente rechazada, por más que esta pueda reproponerse[46].

En el caso del proceso civil peruano se puede identificar claramente esta carga. En efecto, la falta de aportación de medios de prueba por el demandante o demandado que hayan sido mencionados en su acto postulativo podría llevar a que este sea declarada *inadmisible* (o sea, tiene un defecto u omisión, pero puede subsanarse) o, eventualmente, *improcedente* (rechazo por no haberse subsanado). Un juez tiene el *deber* de declarar inadmisible la demanda si se ofrecen medios de prueba, pero no se acompañan a la demanda (según el Código Procesal Civil peruano[47], el ofrecimiento es requisito de la demanda *ex* art. 424 y, a nivel material, los medios de prueba se presentan como anexos de la demanda-documento *ex* art. 425). Así, imaginemos que una parte ofrece un documento simple para acreditar que existe un contrato, pero en el texto de la demanda ha señalado que ofrece una escritura pública. Aquí, el juez, al calificar la demanda, puede ordenar que cumpla con ofrecer el instrumento público, dando oportunidad a la parte para que integre el acervo probatorio. O sea, si la parte quiere ver su demanda admitida, le corresponde presentar la prueba; si no lo hace, no se dará este resultado.

---

[45] Dice la regla: «*If the magistrate judge finds no probable cause to believe an offense has been committed or the defendant committed it, the magistrate judge must dismiss the complaint and discharge the defendant. A discharge does not preclude the government from later prosecuting the defendant for the same offense*». En la doctrina del *common law* se realiza una distinción —no sin cierta imprecisión conceptual— entre *burden of evidence* y *burden of persuasion*. Estas nociones podrían traducirse por «carga de aportación» y «carga de corroboración». Sobre el asunto, desde diversos enfoques, cfr. Kiralfy (1987); Schauer (2009: pp. 223-224); Kaplow (2012); Allen & Stein (2013). La sugerencia de trabajar con las categorías del *common law* ya había sido propuesta por Paula Ramos (2013: pp. 296-297). Para un análisis más puntual, contraponiendo con la realidad del *civil law*, cfr. Nieva Fenoll (2025) y Peixoto (2025), en este último caso analizando críticamente la aplicación de dichos conceptos en la realidad brasileña. Sólo como detalle importante en el ámbito del *common law*, concerniendo la así llamada «carga dinámica de la prueba», se admite la distribución (*shifting*) de la *burden of evidence* (con consecuencias que podrían llegar hasta la pérdida de la causa, pero como sanción por no aportar lo que el juez ordena); pero la segunda jamás se podría alterar (Cfr. McCahey, 2008: p. 8; Nieva Fenoll, 2025).

[46] En este caso trazo un camino diferente a las preocupaciones de Peixoto (2025) respecto de la aplicación del concepto *burden of evidence* al proceso brasileño. Como se aprecia en el cuerpo principal, busco diseñar conceptualmente la categoría para, luego, buscar encontrar algunos ejemplos de ella en el derecho positivo. Más allá de las particularidades del empleo del *burden of evidence* en Estados Unidos (lo cual es bien identificado por Peixoto), esta categoría no debiera ser confundida con la *carga de aportación* que se desarrolla aquí.

[47] Resolución Ministerial N° 010-93-JUS, de 22 de abril de 1993, Texto Único Ordenado del Código Procesal Civil.

La carga de aportación probatoria, pues, al menos en la legislación procesal peruana, se plasma en diversas reglas específicas, pero se consustanciaría en la siguiente proposición normativa genérica: «*Si usted quiere que sus medios de prueba X e Y ingresen formalmente al proceso y que su demanda no sea rechazada, entonces apórtelos*». A diferencia de la tradicional formulación de la «carga de la prueba subjetiva», nótese cómo es que el resultado sí está estrechamente vinculado al comportamiento, siempre que, por supuesto, se cumplan con las reglas procesales que pautean dicho ofrecimiento o aportación.

Aquí podría objetarse que la regla, en realidad, estaría dirigida al juez, colocándole el deber de analizar los requisitos de la demanda y, de constatar alguna omisión o defecto, decidir lo correspondiente. Es incuestionable que existe este deber; sin embargo, esto no quiere decir que no se pueda construir una regla técnica como la mencionada. De hecho, si la parte, a nivel de ofrecimiento de medios probatorios, cumple con todos los requisitos legales, entonces el ordenamiento impone el resultado de que el juez los tenga por formalmente presentados y, por ello, no coloque óbice para la continuación del procedimiento (al menos no por esta razón)[48].

Otra objeción que se puede levantar es que, existiendo el *derecho a la prueba* cuyo contenido normativo ya asegura a las partes el derecho a ofrecer pruebas, a que sean debidamente admitidas y, oportunamente, que sean racionalmente valoradas, no sería necesario ni útil defender otra situación subjetiva, como sería la carga de aportación probatoria al menos tal como aquí es defendida. Para responder a esta crítica, es importante advertir que la existencia de una norma que consagra un derecho subjetivo no necesariamente elimina la figura de la regla técnica. Piénsese en el ejemplo de la televisión antes referido: para encenderla es necesario presionar el botón rojo; pero también existe el derecho de propiedad o posesión sobre el bien mueble que deriva, a su vez, a tener el *derecho a encenderla* (si no hubiese tal derecho habría obstáculos para el disfrute del bien) y el *no deber de no encenderla*. En mi opinión, el derecho subjetivo a la prueba puede coexistir conceptualmente con la carga de aportación si se llega a apreciar que, para la satisfacción del derecho a ofrecer pruebas (dimensión del derecho a la prueba) el juez debe cumplir con su deber de incorporarlas; pero, para ello, la parte, en pleno ejercicio de su esfera de libertad, debe cumplir primero con su carga. Si no se ofrece pruebas, el juez no *debe* incorporarlas, con la correspondiente consecuencia de rechazar la demanda[49].

---

[48] Nótese que a un sistema jurídico puede pertenecer esta norma que expresa la carga de aportación de medios probatorios sin que existan reglas de rechazo liminar. Como se indicó en el texto, el resultado de esta carga en específico es conseguir *incorporar* los medios probatorios al proceso (con cargo a que luego se realice el juicio de admisibilidad); puesto que, si no se ofrecen a tiempo o en la formalidad debida, no ingresarán, con los perjuicios que ello pueda traer. Estos perjuicios pueden ser, por ejemplo, la rebeldía o la imposibilidad de demostrar adecuadamente la hipótesis probatoria.

[49] La carga de aportación probatoria se corresponde íntimamente con la regla que condiciona la alegación de hechos (esenciales) en los actos postulatorios para que se conforme la causa de pedir. Si no se alegan hechos, entonces el juez no podrá pronunciarse sobre ellos ni incorporarlos (principio

Finalmente, ¿cómo vincular la *carga de aportación probatoria* con la corroboración de la hipótesis probatoria[50]? Vinculando con el análisis anterior, coloquemos que la corroboración sería (*B'*). Como es claro, "aportar prueba" (*A)* no garantiza la corroboración ($A \not\subset B'$), por lo que aquella no es una condición suficiente de esta. Por su parte, para que haya corroboración sin duda se requiere de aportación probatoria, pero no necesariamente que una cierta parte la haya aportado ($B' \not\subset A$), por lo que aportar prueba no es una condición necesaria de la corroboración. Imaginemos que *S'* = (aportación de prueba + admisión + peso probatorio), que también es una condición suficiente compleja. Por ello, $S' = A \,\&\, C1 \,\&\, C2$; y $S' \subset B'$; por tanto, $S' \subset A$. Con ello, la *carga de aportación de prueba* no es equivalente a la «carga subjetiva» de la que se habla la doctrina tradicional, y nada más sería que aquello que le corresponde cumplir a una cierta parte para *activar* la condición contribuyente (*A)* para buscar configurar *S'*; esto es, la corroboración probatoria.

6. CONCLUSIONES

Como conclusiones de este trabajo, el concepto tradicional de «carga de la prueba» carece de rigor conceptual y, por tanto, no es útil para dar cuenta de los fenómenos de la decisión en caso de falta de pruebas ni tampoco de los estímulos o incentivos dirigidos a las partes. No es, por tanto, una «carga». Mantener esta terminología no hace más que perpetuar una severa confusión entre la regla de juicio (dirigidas al juez) y las reglas de conducta (dirigidas a las partes).

Por ello, desechando el concepto aún más equívoco de «carga imperfecta», propongo que la categoría lógica de «condición contribuyente» —según la explicación de Von Wright— bien puede ser más adecuada para explicar este estímulo o incentivo probatorio. Es importante notar que este fenómeno existe y, si bien el concepto de «carga» no se muestra como suficiente, sí que resulta necesario dar cuenta de aquel. Negar la utilidad del concepto de carga, tal como se ha hecho, no debiera llevar a

---

dispositivo), lo cual se desvincula con la inexistencia de un deber jurídico de las partes de decir la verdad, entendiéndose este como el deber de alegar *todos* los hechos que, según su conocimiento, habrían ocurrido en la realidad. Pero no sólo ello: como busqué defender en otro lugar (Cavani, 2018 y 2023), la regla de juicio no podría ser alterada por el juzgador, pero, excepcionalmente, y siempre que la parte lo pida, la *carga de aportación* sí podría sufrir una alteración o *shifting* (similar a lo que ocurre en Estados Unidos). Ello ocurriría cuando se constate que la prueba que se necesita para completar alguna no está en posesión de la parte que alegó el hecho al cual dicha prueba está vinculado. Esto, inclusive, tal como defendí, eliminaría la posibilidad de tener que ordenar una prueba de oficio al modo de diligencias para mejor resolver. Por supuesto, la completitud de la laguna probatoria a partir de esta alteración podría generar la imposición de deberes probatorios en las partes, e inclusive sanciones, en una propuesta de alguna manera cercana a la de Paula Ramos (2020). Dejo estas reflexiones para otra oportunidad.

[50] Aquí, por supuesto, hablo del ámbito de la valoración de la prueba, y no de la suficiencia de la hipótesis ya corroborada, que tiene que ver con los estándares de prueba y la justificación de la decisión probatoria. Por todos, Ferrer Beltrán (2025).

negar dicha realidad. Así, en efecto, se trata de un *réquiem*, pero, dado que no todo está muerto, también es una *obertura*.

A partir de ello, con la condición contribuyente queda claro que el aportar prueba no es una condición suficiente para vencer en la causa, pero sí es una propiedad que necesariamente debe estar presente, dentro de una cadena compleja de condiciones, que permiten alcanzar una sentencia favorable. Este cambio de enfoque permite explicar la estrategia procesal sin forzar categorías dogmáticas que no parecen conducir a ningún lugar.

A diferencia de la noción general de «carga de la prueba», pienso que sí existe una auténtica carga procesal cuya autonomía conceptual es posible defender: la *carga de aportación probatoria*. He buscado demostrar que existe una relación causal directa y dependiente de la libertad de la parte: si se cumple con aportar los medios probatorios conforme a la norma, se obtiene el resultado normativo de su admisión formal y se evita el rechazo liminar (inadmisibilidad o improcedencia) de la demanda.

Finalmente, el esclarecimiento conceptual aquí propuesto está lejos de agotar el debate; quisiera pensar, más bien, que contribuye a sentar las bases para una futura reconstrucción dogmática. En efecto, al depurar el concepto de «carga de la prueba» y restringirlo a la estricta «aportación», quedaría el camino libre para determinar cómo trabajar con esta carga, como sería el caso, por ejemplo, de alterarla a fin de promover una mayor completitud del material probatorio sin sacrificar garantías fundamentales de las partes. La carga de aportación probatoria, por tanto, sería aquello que requiere cumplir la parte para activar la condición contribuyente, a saber: el cumplimiento de una condición necesaria de una condición suficiente que, en este caso, genere un resultado positivo: que la hipótesis probatoria sea corroborada.

REFERENCIAS BIBLIOGRÁFICAS

Fuentes doctrinarias

Allen, R. & Stein, A. (2013). Evidence, Probability, and the Burden of Proof. *Arizona Law Review*, *55*, 557-602.

Augenti, G. P. (1932). *L'onere della prova*. Foro Italiano.

Augenti, G. P. (2000). Apéndice [1947]. *La prueba civil*, 2ª ed. (F. Carnelutti). Depalma, 209-256.

Buril de Macedo, L. & Peixoto, R. (2014). *Ônus da prova e sua dinamização. (*2ª ed. 2016). JusPodivm.

Calmon de Passos, J. J. (2009). *Esboço de uma teoria das nulidades aplicadas às nulidades processuais*, 1ª ed., 4° tiraje. Rio de Janeiro.

Carpes, A. (2010). *Ônus dinâmico da prova*. Livraria do Advogado.

Carpes, A. (2015). Carga de la prueba dinámica y el postulado de la razonabilidad. En R. Cavani y V. de Paula Ramos (coords.), *Prueba y proceso judicial* (pp. 349-362). Instituto Pacífico (traducción de R. Cavani).

Carpes, A. (2017). *Ônus da prova no novo CPC. Do estático ao dinâmico*. Revista dos Tribunais.

Carpes, A. (s/f). Dinamização do ônus da prova: do ônus de justificação ao dever de fundamentação (artículo inédito gentilmente cedido por el autor).

Cavani, R. (2014). *La nulidad en el proceso civil.* Palestra.

Cavani, R. (2019). «Prueba de oficio» y «carga de la prueba»: Una propuesta equilibrada. En R. Cavani (coord.), *Garantías procesales y poderes del juez* (pp. 277-303). Zela.

Cavani, R. (2023). «Prueba de oficio» y «carga de la prueba»: Una propuesta equilibrada. En R. Cavani (auth.); S. Espejo (ed.), *El proceso civil en su laberinto: Ensayos para una reforma* (pp. 509-532). Zela.

Comoglio, L. P. (2013). Art. 2697 - Onere della prova. En G. Alpa y V. Mariconda (eds.), *Commentario del Codice Civile* (pp. 291-335). Wolters Kluwer.

Conso, G. (1969). *Istituzioni di diritto processuale penale*, 3ª ed. Giuffrè.

Costa, E. J. F. (2017). Presunção de inocência civil: algumas reflexões no contexto brasileiro. *Revista brasileira de direito processual*, *100,* 129-144.

Didier Jr., F. (2017). A distribuição legal, jurisdicional e convencional do ônus da prova no novo Código de Processo Civil brasileiro. *Revista de Direito Mackenzie*, *11*(2), 129-155.

Didier Jr., F. & Nogueira, P. H. P. (2012). *Teoria dos fatos jurídicos processuais.* JusPodivm.

Falcón, E. M. (2003). *Tratado sobre la prueba (civil, comercial, laboral, penal, administrativa)*, tomo 1. Astrea.

Fernández López, M. (2006). *La carga de la prueba en la práctica judicial civil.* La Ley.

Ferrer Beltrán, J. (2005). *Prueba y verdad en el derecho*, 2ª ed. Marcial Pons.

Ferrer Beltrán, J. (2019). La carga dinámica de la prueba. Entre la confusión y lo innecesario. En J. Nieva Fenoll, J. Ferrer Beltrán y L. Giannini, *Contra la carga de la prueba* (pp. 53-87). Marcial Pons.

Ferrer Beltrán, J. (2025). *Prueba sin convicción: Estándares de prueba y debido proceso*, 2ª ed. Marcial Pons.

Gavazzi, G. (1970). *L'onere: tra libertà e obbligo.* Giappichelli.

Hohfeld, W. N. (2004). *Conceptos jurídicos fundamentales*, 1ª ed., 5ª reimp. [1913]. Fontamara.

Kaplow, L. (2012). Burden of Proof. *The Yale Law Journal*, *121*(4), 738-859.

Kiralfy, A. (1987). *The Burden of Proof.* Butterworths Law.

Mitidiero, D. (2020). O ônus da prova e seus inimigos. *Revista de Processo*, *306,* 17-47.

McCahey, J. (2008). The Burdens of Production and Persuasion. *Proof. The Journal of the Trial Evidence Committee*, *16*(3), 7-10.

Micheli, G. A. (1966). *L'onere della prova* [1942]. Cedam.

Nieva Fenoll, J. (2019). La carga de la prueba: una reliquia histórica que debiera ser abolida. En J. Nieva Fenoll, J. Ferrer Beltrán y L. Giannini, *Contra la carga de la prueba* (pp. 23-52). Marcial Pons.

Nieva Fenoll, J. (2024). *Trilogía sobre la carga de la prueba: hacia su abolición.* Palestra.

Nieva Fenoll, J. (2025). Autopsia de la carga de la prueba. *InDret*, 3.2025, 430-450.

Paula Ramos, V. (2013). Derecho fundamental a la prueba. *Gaceta Constitucional*, *65*, 286-299 (trad. R. Cavani).

Paula Ramos, V. (2015). Ônus da prova no processo civil. Do ônus ao dever de provar. Revista dos Tribunais.

Paula Ramos, V. (2015a). Cargas y deberes probatorios de las partes en el nuevo CPC brasileño. En R. Cavani y V. de Paula Ramos (coords.), *Prueba y proceso judicial* (pp. 363-382). Instituto Pacífico (trad. R. Cavani).

Paula Ramos, V. (2020). *La carga de la prueba en el proceso civil: de la carga al deber de probar.* Marcial Pons (trad. L. Criado Sánchez).

Paula Ramos, V. (2021). Novos debates sobre o «ônus» da prova: acordos e desacordos entre a doutrina sobre o tema. *Revista Acadêmica da Faculdade de Direito do Recife*, 247-262.

Peixoto, R. (2025). A bipartição do ônus da prova no common law: análise comparativa e implicações para o direito processual brasileiro. *Revista de Processo*, *368,* 305-327.

Peyrano, J. W. (2004). Doctrina de las cargas probatorias dinámicas y la máquina de impedir en materia jurídica. En J. W. Peyrano (dir.), I. Lepori White (coord.), *Cargas probatorias dinámicas* (pp. 75-98). Rubinzal-Culzoni.

Prütting, H. (2010). Carga de la prueba y estándar probatorio: la influencia de Leo Rosenberg y Karl Heinz Schwab para el desarrollo del moderno derecho probatorio. *Revista Ius et Praxis*, *1*, 453-464 (trad. A. Pérez Ragone).

Rivera Morales, R. (2010). *La prueba: un análisis racional y práctico.* Marcial Pons.

Rodríguez Álvarez, A. (2020). *La carga de la prueba en supuestos de discriminación: Su regulación en el proceso civil.* Tirant lo Blanch.
Rosenberg, L. (2002). *La carga de la prueba* [1956]. EJEA (trad. E. Krotoschin).
Schauer, Frederick. (2009). *Thinking like a Lawyer. A New Introduction to Legal Reasoning.* Harvard University Press.
Verde, G. (1974). *L'onere della prova ne processo civile.* Jovene.
Vogt Geisse, T. (2025). La carga de la prueba como fundamento del esclarecimiento de los hechos y del manejo de la confidencialidad en el proceso civil. *Latin American Legal Studies, 13*(2)*,* 193-231.
Von Wright, G. H. (1963). *Norm and Action: A Logical Inquiry.* Routledge & Kegan Paul.
Von Wright, G. H. (2001). *A Treatise on Induction and Probability* [1951]. Routledge.

Fuentes legislativas

Perú. Decreto Legislativo N° 295, de 25 de julio de 1984, Código Civil. *https://spij.minjus.gob.pe/spij-ext-web/#/detallenorma/H682684*
Perú. Resolución Ministerial N° 010-93-JUS, de 22 de abril de 1993, Texto Único Ordenado del Código Procesal Civil. *https://spij.minjus.gob.pe/spij-ext-web/#/detallenorma/H682685*
Perú. Decreto Legislativo N° 957, de 29 de julio de 2004, Código Procesal Penal. *https://spij.minjus.gob.pe/spij-ext-web/#/detallenorma/H682695*
Perú. Ley N° 29497, de 15 de enero de 2010, Ley Procesal del Trabajo. *https://spij.minjus.gob.pe/spij-ext-web/#/detallenorma/H1001319*

Quaestio facti. Revista Internacional sobre Razonamiento Probatorio
Quaestio facti. International Journal on Evidential Legal Reasoning
Sección: Ensayos
2026 I 11 pp. 63-88
Madrid, 2026
DOI: 10.33115/udg_bib/qf.i11.23248
Marcial Pons Ediciones Jurídicas y Sociales
© Sebastián Bravo Ibarra
© Eduardo Estrada Ravena
ISSN: 2604-6202
Recibido 17/02/26 | Aceptado 28/05/26 | Publicado *online*: 08/06/2026
Editado bajo licencia Reconocimiento 4.0 Internacional de Creative Commons

MÁS ALLÁ DE LA INTERFAZ: DESLINDES Y SINGULARIDADES DEL *TRUTHMAKER* DIGITAL

Sebastián Bravo Ibarra*

Eduardo Estrada Aravena**

RESUMEN: Este trabajo examina la función del hecho digital como *truthmaker* en el razonamiento probatorio. Sostiene que los objetos digitales que se presentan como prueba en el proceso —capturas de pantalla, fotografías digitales, registros, *logs* u otras visualizaciones— son formas de acceso a un tipo de hecho cuya materialidad es distribuida, codificada y relacional, que no pueden ser conocidos adecuadamente mediante los mecanismos tradicionales de percepción o inspección, limitados a la superficie de ese fenómeno. Sobre esta base, el artículo desarrolla las singularidades estructurales de los hechos digitales y muestra que ellas exigen desplazar la valoración probatoria desde una epistemología de la presencia, propia del paradigma analógico, a una epistemología de la producción técnica. Finalmente, analiza las consecuencias que este giro conlleva para la teoría de la prueba.

PALABRAS CLAVE: *truthmaker*; prueba digital; hechos digitales; razonamiento probatorio; decisión judicial

* Abogado. Licenciado en Derecho por la Pontificia Universidad Católica de Chile. Doctor en Derecho por la Universidad de Chile. Master in Global Rule of Law and Constitutional Democracy por la Università degli Studi di Genova. Máster en Razonamiento Probatorio por la Universitat de Girona. Juez civil y laboral en el Primer Juzgado de Letras de Rengo.

** Abogado. Licenciado en Ciencias Jurídicas y Sociales por la Universidad de Chile. Máster en Derecho Probatorio en el Proceso Penal por la Universidad de Barcelona. Máster en Argumentación Jurídica por la Universidad de León. Relator de la Iltma. Corte de Apelaciones de Santiago.

BEYOND THE INTERFACE: DEMARCATIONS AND SINGULARITIES OF THE DIGITAL TRUTHMAKER

ABSTRACT: This article examines the function of the digital fact as a truthmaker in evidentiary reasoning. It argues that the digital objects submitted as evidence in judicial proceedings—screenshots, digital photographs, records, logs, or other visualizations—are forms of access to a type of fact whose materiality is distributed, encoded, and relational, and which cannot be adequately known through traditional mechanisms of perception or inspection, as these remain confined to the surface of the phenomenon. On this basis, the article develops the structural singularities of digital facts and shows that they require shifting evidentiary assessment from an epistemology of presence, characteristic of the analog paradigm, toward an epistemology of technical production. Finally, it analyzes the consequences of this shift for the theory of evidence.

KEYWORDS: truthmaker; digital evidence; digital facts; evidential reasoning; judicial decision.

SUMARIO: 1. INTRODUCCIÓN.— 2. UNA RECAPITULACIÓN ONTOLÓGICA.— 3. SINGULARIDADES DEL HECHO DIGITAL: 3.1. Estructurabilidad técnica; 3.2. Registrabilidad originaria; 3.3. Alterabilidad; 3.4. Ubicuidad; 3.5. Temporalidad difusa: 3.6. Una singularidad entre singularidades: estratificación en un sistema de planos o multiplanar.— 4. DESLINDES CON LAS CATEGORÍAS PROBATORIAS TRADICIONALES: 4.1. El hecho técnico y sus límites explicativos; 4.2. El hecho institucional y sus límites explicativos; 4.3. El hecho digital como categoría ontológica autónoma.— 5. DE LA EPISTEMOLOGÍA DE LA PRESENCIA A LA EPISTEMOLOGÍA DE LA PRODUCCIÓN TÉCNICA: 5.1. Tres regímenes de comparecencia del hecho; 5.2. De la correspondencia a la producción.— 6. CONSECUENCIAS PARA LA TEORÍA DE LA PRUEBA: 6.1. El colapso del paradigma del cotejo; 6.2. La crisis del modelo perceptivo de valoración; 6.3. La insuficiencia de la prueba pericial; 6.4. La verificación arquitectónica a partir del truthmaker digital; .— 7. CONCLUSIONES.— BIBLIOGRAFÍA.

1. INTRODUCCIÓN

La expansión masiva de tecnologías digitales y, más recientemente, la irrupción de sistemas de inteligencia artificial generativa han alterado de manera estructural las condiciones de producción, circulación y registro de los hechos relevantes para el derecho. En el ámbito probatorio, esta transformación no se limita a la simple aparición de nuevos soportes o medios de prueba, sino que compromete los presupuestos epistemológicos sobre los cuales se ha construido históricamente el conocimiento judicial de los hechos. El derecho probatorio moderno —forjado en torno a la percepción sensorial y la inmediación— se enfrenta hoy a fenómenos que ya no pueden ser comprendidos adecuadamente bajo esos presupuestos[1].

[1] En el número anterior de Quaestio facti, Amalia Amaya y Andrés Páez, al reflexionar sobre el legado de Larry Laudan y el futuro de la epistemología jurídica, ya advertían que la irrupción de la inteligencia artificial generativa, el uso desprevenido de grandes modelos de lenguaje por operadores

Los objetos digitales que se presentan como prueba en el proceso —publicaciones en redes sociales, registros de acceso, metadatos, *logs* de sistemas, decisiones automatizadas o contenidos sintéticos generados por inteligencia artificial— ya no pueden entenderse como una mera adaptación de la prueba tradicional ni tampoco pueden homologarse, sin más, a la noción de documento electrónico[2]. Tales concepciones no hacen más que utilizar categorías dogmáticas decimonónicas para hacer frente a los problemas del presente y del futuro. Los objetos digitales son más bien formas de acceso a hechos que ocurren en un dominio distinto al de los hechos analógicos, en la medida que los procesos que los constituyen tienen lugar en arquitecturas técnicas específicas. Así, una captura de pantalla de un mensaje no es el envío de ese mensaje, sino un objeto digital derivado que presenta, bajo una forma perceptible para un observador, ciertos efectos de una operación técnica más compleja.

Esta distinción tiene relevancia para la teoría de la prueba en la medida que modifica el modo en que una proposición puede ser considerada verdadera o falsa cuando se refiere a un hecho digital. De acuerdo con una teoría correspondentista de la verdad, resulta fundamental considerar la relación existente entre el lenguaje y el mundo o, más precisamente, entre las proposiciones (*truthbearers*) y los hechos que las confirman o refutan (*truthmakers*). Mientras el *truthbearer* es una proposición que aspira a la corrección veritativa, el *truthmaker* es la realidad extralingüística que hace verdadera o falsa esa proposición. En este sentido, las partes aportan una serie de proposiciones acerca de los hechos que, en último término, permiten al juez construir un *thema probandum* en concreto, para luego determinar, a través de un *iter* argumentativo, si existen razones justificativas para su confirmación o refutación en cuanto *truthbearer*.

Este esquema supone, sin embargo, un cierto tipo de *truthmaker* analógico, que se diferencia ontológicamente de lo que aquí denominaremos *truthmaker* digital. Bajo el paradigma analógico, el hecho que hace verdadera una proposición suele con-

---

jurídicos y la proliferación de contenidos sintéticos —como los *deepfakes*— podrían generar una crisis significativa en la valoración de la prueba, exigiendo una reevaluación del peso y del riesgo epistémico de diversos tipos de evidencia. Véase Amaya y Páez (2026), p. 27.

[2] La mayor parte de la doctrina sostiene que una prueba de estas características se encuentra en el mismo plano de la prueba documental. Por ejemplo, Acosta-León, C. (2023), p. 10, sostiene que «estos documentos digitales o electrónicos, en términos tecnológicos, están almacenados en un tipo de objeto digital denominado archivo o fichero que sirve de contenedor de diversos tipos de documentos binarios. Dependiendo de su contenido, son de varios tipos: archivos de audio, sonido o voz, archivos de video […], archivos de texto, archivos de imágenes, fotos o gráficos, archivos de programas fuentes y ejecutables, archivos comprimidos, archivos de mensajes de correo o chat, archivos de configuración, archivos de transacciones bancarias, archivos de imagen de disco, datos de audio y video en flujo continuo o en *streaming*, etcétera». Uno de los errores principales de estos autores es simplificar, homologar y subsumir toda clase de información a la categoría de documento, por el solo hecho de ser susceptibles de almacenamiento en un archivo, poniendo toda su atención en el continente, y olvidando el contenido. Por otro lado, hay autores que han advertido este problema y han instado por una regulación de la prueba digital como un medio de prueba autónomo. En este sentido, cfr. Santos, A. (2025), pp. 10 y ss.; y Galvis, A. y Bustamante, M. (2019), p. 192.

cebirse como un evento que se verifica y percibe en el mundo analógico, i.e. en un espacio determinado, en una temporalidad relativamente lineal y presentado ante un «observador», mediante percepción directa o inspección técnica. Se trata de un presupuesto que rara vez resulta problemático pues se sostiene en la relativa estabilidad del mundo material. Sin embargo, con la irrupción de la prueba digital esa relación se tensiona, en la medida que el hecho digital hacia el cual apuntan, presenta una serie de singularidades que rompen con esa estabilidad.

En el contexto antes señalado, este artículo se inscribe en una línea de investigación abierta por uno de los autores en un trabajo anterior[3], donde se argumentó acerca de la necesidad de distinguir las diferencias entre lo analógico y lo digital, a un nivel no simplemente superficial, sino ontológico; y donde sugiere, además, una metodología abductivo-contractiva para la valoración de la prueba digital. Aquellas tesis se asumen aquí como punto de partida para abordar dos cuestiones específicas: ¿qué tipo de *truthmaker* es un hecho digital? Y, respondido lo anterior, ¿qué consecuencia tiene para la teoría de la prueba que el verificador pertenezca a un régimen ontológico distinto al que se presupone?

El recorrido se desarrolla en cuatro movimientos. En primer lugar, se desarrolla un análisis estructural del hecho digital que profundiza acerca de seis singularidades, cuya coherencia constitutiva define su modo de existencia. En segundo lugar, se sostiene que, una vez caracterizado en estos términos el hecho digital, no se deja subsumir en las categorías clásicas del hecho técnico ni del hecho institucional, y debe ser reconocido como un *truthmaker* autónomo inescindible de sus singularidades. En tercer lugar, se argumenta que todo lo anterior exige abandonar lo que aquí denominamos «epistemología de la presencia» heredada del paradigma analógico y reconocer que el hecho digital no se presenta ni se representa, sino que es producido por una arquitectura técnica como un resultado renderizado, respondiendo más bien a lo que llamamos «epistemología de la producción técnica». Finalmente, se derivan las consecuencias estructurales que estos hallazgos imponen sobre los pilares dogmáticos de la teoría de la prueba.

2. UNA RECAPITULACIÓN ONTOLÓGICA

Tradicionalmente, la doctrina y el proceso judicial han operado con un paradigma analógico de los hechos, entendidos como eventos que ocurren en el mundo material, susceptibles de reconstrucción mediante los distintos medios de prueba. La irrupción de contenidos sintéticos capaces de emular la realidad ha puesto en crisis ese modelo, en la medida que aquello que se presenta ante el juez —o evaluador,

[3] Si bien el artículo Estrada (2025) cimentó la formulación ontológica del concepto de hecho digital y sus consecuencias para la valoración probatoria, el desarrollo sistemático del problema requería incorporar herramientas propias de la filosofía analítica y la epistemología, lo que explica un desarrollo en extenso, en un nivel de abstracción diferente, a través del presente trabajo en coautoría.

en su caso— como imagen, audio, video o registro, puede no tener como correlato un acontecimiento material previo, sino haber sido producido digitalmente. Sin embargo, como fuera desarrollado por Estrada recientemente, fenómenos como los *deepfakes* solo vienen a radicalizar una deuda conceptual previa: la falta de desarrollo de una teoría que aborde la naturaleza de los hechos digitales y la forma en que deben ser abordados por la teoría de la prueba[4].

En el trabajo antes mencionado, el hecho digital fue definido desde la ontología de la individuación técnica desarrollada por Gilbert Simondon y la caracterización del modo de existencia de los objetos digitales propuesta por Yuk Hui[5], como un «acontecimiento informacional que consiste en una operación o cambio de estado producido dentro de un sistema técnico, cuya arquitectura lo inscribe e identifica como una misma entidad conforme a sus reglas internas, con independencia de que exista o no un correlato material en el mundo físico»[6], de la cual destacamos aquí ese cambio de estado o, si se quiere, una «transición de estado» producida dentro de una arquitectura informacional determinada.

Pues bien, para una mayor claridad expositiva, comenzaremos diferenciando los hechos digitales de los objetos digitales. Para Hui, una fotografía digital almacenada en un teléfono no es solamente la imagen que el usuario ve en la pantalla cuando abre el archivo. Corresponde a un objeto digital compuesto por una estructura de datos, un formato de codificación, metadatos y relaciones con otras aplicaciones o servicios que permiten abrirla, copiarla, compartirla, editarla o eliminarla. Su existencia no se limita a la percepción visual de la imagen sino que se extiende al conjunto de relaciones lógicas y técnicas que la componen, y que permiten al sistema reconocerla como el mismo objeto a través de distintas operaciones.

El hecho digital se sitúa en otro plano. A partir de la idea de objetos digitales de Hui, el concepto pretende hacer referencia a la variación que ocurre sobre esos objetos digitales dentro de una arquitectura informacional. Mientras el objeto digital designa una configuración relativamente estabilizada —la fotografía como archivo reconocible por el sistema—, el hecho digital designa la operación que crea, modifica, transmite, elimina o reconfigura esa configuración. Tomar una fotografía con el teléfono, editarla, enviarla por una aplicación de mensajería, subirla a una red social o eliminarla del dispositivo son hechos digitales en la medida que producen transiciones de estado dentro del sistema. En un sentido más técnico aun, existen hechos específicos del mundo digital, que inciden en objetos digitales también específicos, tales como la rasterización, la vectorización, la transcodificación, la multiplexación y demultiplexación, la parserización, entre otros, que no conocemos por los sentidos sino solamente a través de sus objetos finales renderizados.

---

[4] Estrada (2025), p. 125.

[5] Simondon (2007), pp. 56-82; Hui (2012), p. 387; Hui (2016), p. 111. De manera análoga al ejercicio que Hui realiza con la teoría de la individuación de Simondon aplicada a los objetos digitales, extendemos aquí ese ejercicio a la dimensión fáctica, de interés para la teoría de la prueba.

[6] Estrada (2025), p. 131.

Se trata de una distinción que permite sortear una confusión habitual. La captura de pantalla de un mensaje puede ser el medio por el cual se introduce al proceso cierta información, pero no necesariamente la fuente más adecuada para reconstruir el hecho digital consistente en el envío, recepción o visualización de ese mensaje. Del mismo modo, un reporte exportado desde un sistema puede ser un objeto digital relevante, pero su valor probatorio dependerá de su relación con la arquitectura que produjo el registro, con las reglas técnicas que lo estabilizan y con las operaciones que pretende representar.

La definición antes propuesta desplaza el análisis desde una lógica representacional hacia una lógica operativa. El hecho digital no es el dato visible, el archivo almacenado ni la interfaz que lo exhibe, sino la transición misma que altera el estado del sistema conforme a reglas técnicas específicas. No se trata, en definitiva, de un evento del mundo material que luego es fijado en un soporte, sino de una modificación estructural cuya existencia coincide con su inscripción en la arquitectura que la produce.

Ahora bien, la inscripción originaria no garantiza la permanencia indefinida del registro. Un archivo puede ser eliminado, una entrada puede ser sobrescrita, un sistema puede ser reconfigurado. Sin embargo, la eventual supresión o alteración posterior no elimina el carácter transicional del hecho digital: lo que puede desaparecer es el acceso al estado resultante o a sus huellas, no la transición que tuvo lugar dentro de la arquitectura en un momento determinado. Esta distinción resulta decisiva para el razonamiento probatorio, pues obliga a distinguir entre la existencia del hecho digital y la disponibilidad actual de su registro.

Así, como ocurre en una de las figuras paradigmáticas del cibercrimen —el acceso indebido o la alteración no autorizada de datos—, el *hacker* que ingresa a un sistema sin credenciales válidas no «rompe» el código ni interrumpe su lógica interna; lo que hace es eludir las condiciones de autorización previstas por la arquitectura. El sistema continúa ejecutando operaciones conforme a sus reglas técnicas.

3. SINGULARIDADES DEL HECHO DIGITAL

Los hechos digitales poseen un conjunto de propiedades estructurales que singularizan su modo de existencia, y que resultan determinantes para cualquier pretensión de valoración racional en sede judicial, entre las que destacan las siguientes: estructurabilidad técnica, registrabilidad originaria, alterabilidad, ubicuidad, temporalidad difusa y estratificación en un sistema de planos o multiplanar.

A continuación, examinaremos estas seis singularidades con el propósito de precisar en qué sentido los hechos digitales difieren concretamente de los hechos analógicos y por qué dicha diferencia exige repensar los fundamentos epistemológicos del derecho probatorio, al conformar un auténtico *truthmaker* digital[7].

[7] Estas singularidades suelen ser connaturales a los hechos digitales, pero no necesariamente los definen en un sentido lógico. Nótese también que rasgos de esta naturaleza suelen atribuirse a la

3.1. Estructurabilidad técnica

La primera singularidad estructural del hecho digital es su estructurabilidad técnica. El hecho digital no puede existir sin la arquitectura que lo produce, más bien emerge de la estructura que define las condiciones de su existencia, inscripción y validación. Se trata, en definitiva, de un *truthmaker* que no emerge espontáneamente en el mundo material sino que se produce dentro de una arquitectura previamente diseñada, aunque su ocurrencia concreta pueda ser intencional, automática o derivada de la interacción de sistemas.

A diferencia del hecho analógico, cuya ocurrencia es independiente de las reglas que posteriormente lo describen o califican, el hecho digital solo existe en la medida en que se produce conforme a una arquitectura técnica que define previamente qué cuenta como evento, qué transiciones de estado son posibles, cómo se registra e integra en el sistema, bajo qué estados puede persistir y qué condiciones permiten reconocerlo como válido dentro del sistema. No se trata simplemente de que el hecho digital «utilice» tecnología para ser constatado sino que «necesita» de dicha tecnología para existir del modo en que lo hace.

Este rasgo explica por qué yerra la doctrina al tratar estos fenómenos como si fueran meramente «cuestiones técnicas» delegables a la prueba pericial en caso de controversia, como si se tratase solamente de «llamar a alguien» que tenga un *know-how* en la utilización de dispositivos tecnológicos[8]. Reducir el problema a la intervención de expertos supone confundir la estructuración ontológica del hecho con su mera constatación empírica.

3.2. Registrabilidad originaria

La segunda singularidad estructural del hecho digital es su inscripción o registrabilidad originaria. En la medida que el hecho digital consiste en una transición de estado producida al interior de una arquitectura informacional determinada, su existencia implica alguna forma de inscripción operativa en sistema, aunque no necesariamente un registro persistente, visible o recuperable. No existe un momento previo al registro ni una exterioridad desde la cual el hecho pueda luego ser fijado. Así las cosas, el *truthmaker* digital tiene la ventaja de verificar de una forma mucho más ostensible que el *truthmaker* analógico, precisamente en virtud de su inscrip-

---

evidencia digital, en general. Así, por ejemplo, Brighi y Ferrazzano (2021), p. 16, incluyen dentro de los atributos de la prueba digital los siguientes: intangibilidad, alterabilidad, cambios debidos al uso regular o mal manejo, volatilidad y reproductibilidad potencialmente ilimitada. Sin embargo, estos atributos ponen el foco en el medio de prueba utilizado como información en vez del hecho digital que sirve como *truthmaker* de aquellas proposiciones que son objeto de prueba.

[8] *Cfr*. Abel Lluch y Picó i Junoy. (2011), p. 151 y ss.

ción *ab initio*. El *truthmaker* analógico no implica su conservación ni su trazabilidad posterior. Un suceso puede acontecer y desaparecer sin dejar vestigios estructurales suficientes para su reconstrucción

La registrabilidad, más allá de la simple acumulación de datos, es una consecuencia necesaria de la ejecución técnica de la operación que constituye el hecho. Toda transición digital implica una modificación estructural del sistema: actualización de una base de datos, escritura en memoria, generación de *logs*, alteración de parámetros o activación de registros internos. El hecho digital consiste precisamente en esa modificación inscrita conforme a reglas técnicas determinadas.

Ahora bien, todo esto deja huellas dentro del sistema, pero que eventualmente podrían ser modificadas o suprimidas mediante nuevas transiciones, dificultando finalmente su acceso. Debe distinguirse, entonces, entre inscripción, permanencia y accesibilidad del *truthmaker digital*. La registrabilidad originaria no significa que su registro sea inmutable ni indefinidamente accesible. Podría, en efecto, eliminarse o alterarse en un caso concreto, pero al hacerlo se generaría una nueva transición de estado, también inscrita, que contribuiría al análisis para la verificación del *truthbearer*.

3.3. Alterabilidad

Una tercera singularidad de los hechos digitales corresponde precisamente a su alterabilidad técnica. A diferencia de los hechos analógicos, que no pueden ser modificados en su materialidad una vez ocurridos, los hechos digitales pueden ser alterados, reconfigurados o actualizados dentro del propio entorno técnico que los produce y conserva. Esta alterabilidad no opera únicamente en el plano del relato o de la interpretación posterior, sino que incide directamente en la existencia informacional del hecho, afectando su contenido, su forma de inscripción y las condiciones técnicas bajo las cuales puede ser accedido y reutilizado. Además, tal como sostiene Bruno Latour, no sólo los hechos sino también las máquinas cambian constantemente, de manera que no nos encontramos frente a meras reproducciones estáticas, sino ante ensamblajes dinámicos[9].

Desde una perspectiva probatoria, esta característica obliga a revisar críticamente los conceptos tradicionales de integridad e inmutabilidad. En el modelo clásico, la integridad de la prueba se asocia a la conservación material del objeto tal como fue producido, de modo que cualquier alteración posterior suele interpretarse como un indicio de falsedad o pérdida de valor epistémico. Así, quienes teorizan sobre la evidencia digital en materia penal, generalmente recurren a la noción de cadena de custodia, como mecanismo destinado a garantizar que el medio de prueba no ha sido intervenido, sustituido o modificado desde su obtención hasta su incorporación y valoración en el proceso, asegurando así la continuidad material del soporte y la fiabilidad de su contenido.

[9] Latour (1987), p. 134.

Sin embargo, en el plano digital, esta presunción resulta insuficiente y engañosa[10]. La posibilidad de modificación no constituye una anomalía, sino una condición ordinaria del funcionamiento de los sistemas informáticos. En este contexto, la confianza acrítica en herramientas como el *hash* —concebido a veces como un equivalente funcional de la inmutabilidad— puede inducir a una falsa sensación de certeza. El *hash* permite verificar que un archivo no ha cambiado desde un momento determinado, pero nada dice sobre sus condiciones de producción, sobre eventuales modificaciones previas a su fijación, ni sobre la fiabilidad del sistema que lo generó, almacenó o presentó.

En este escenario, un archivo generado íntegramente por inteligencia artificial puede mantener perfecta estabilidad criptográfica desde el momento de su creación y, aun así, carecer de fiabilidad probatoria respecto de la proposición fáctica que pretende acreditar. Es así como el *truthmaker* digital se diferencia una vez más del *truthmaker* analógico en cuanto que, en el primero, la alteración del objeto que permite acceder a él forma parte de sus propiedades esenciales, sin que por ello su fiabilidad quede automáticamente destruida; mientras que, en el segundo, la alteración del objeto que permite acceder a él, es indiciaria de poca fiabilidad. En este último caso, existirán razones justificativas que apunten hacia la falsedad del *truthbearer*, mientras que en el primero no. El problema, entonces, no es simplemente si el archivo fue alterado, sino si la arquitectura técnica que produjo esa transición de estado permite realmente construir argumentos en apoyo o no de su fiabilidad.

3.4. Ubicuidad

Otro rasgo estructural de los hechos digitales es su ubicuidad. A diferencia del hecho analógico, que ocurre en un lugar determinado del espacio físico, el hecho digital no se circunscribe a un único punto espacial singular. Ello no significa que carezca de soporte material, sino que su ocurrencia no se agota en la localización física de un dispositivo singular. En este sentido, Ingvarsson señala que lo digital tiende a desanclarse del esquema clásico de localización espacial que presupone una correspondencia directa entre acontecimiento y lugar, ya que tenemos sistemas ubicuos donde se distribuye información y ya no artefactos[11].

La transición de estado que constituye el hecho digital puede involucrar múltiples componentes de una arquitectura distribuida: dispositivos de origen, servidores intermedios, sistemas de respaldo o nodos replicados en distintas jurisdicciones. En este sentido, el hecho digital no coincide plenamente con ninguno de sus soportes físicos particulares, sino con la modificación estructural del sistema en el que se integra funcionalmente.

---

[10] Cardona (2020), pp. 46-49 plantea, a propósito de la prueba digital, el carácter esencialmente volátil, eliminable, alterable y modificable.

[11] Ingvarsson, J. (2021), p. 88.

Esta ubicuidad es una propiedad ontológica que va más allá de la mera facilidad de circulación o copia. Un mismo evento informacional puede manifestarse concurrentemente en distintos dispositivos, servidores, interfaces y bases de datos, sin que ninguna de esas instancias pueda, por sí sola, agotar el «lugar» donde el hecho verdaderamente ocurre. Lo cierto es que el hecho digital no está contenido en esos soportes, sino que se actualiza en ellos conforme a reglas técnicas[12] de acceso y sincronización.

A nivel forense, Mason ilustra el punto con el siguiente ejemplo: un pago realizado con tarjeta de débito en una cafetería de Ayrshire (Escocia) fue registrado por el sistema bancario como realizado en una cafetería homónima de Lynmouth a más de 700 kilómetros de distancia. Más allá de identificar en qué lugar se produce el error —aunque un perito podría determinar que el sistema funcionó correctamente— el autor destaca que un sistema técnico puede operar formalmente en distintos planos y lugares de validación, de modo que, muchas veces, será de poco interés probatorio lo que el sistema muestra en su interfaz[13].

Desde el punto de vista jurídico, esta característica introduce tensiones relevantes respecto de los criterios tradicionales de competencia territorial. Si el hecho analógico permite identificar con relativa claridad un *locus commissi*, el hecho digital puede involucrar infraestructuras situadas en distintas jurisdicciones, dificultando su adscripción a un único espacio procesal. En consecuencia, la ubicuidad del hecho digital no significa que este «ocurra en todas partes y en ninguna a la vez», sino que su modo de existencia no se encuentra anclado a un punto espacial singular; más bien se distribuye conforme a la lógica técnica de la arquitectura que lo produce.

Como adelantamos, la ubicuidad del hecho digital tiene consecuencias epistémicas relevantes. Por una parte, la verificación mediante un *truthmaker* ubicuo nos aleja de la idea tradicional de identificar un original único con el cual cotejar copias, pues pueden coexistir múltiples registros primarios del mismo evento. Por otra, desplaza el foco del análisis desde el soporte hacia el sistema: lo relevante ya no es dónde «está» el hecho, sino bajo qué condiciones técnicas se produce, replica y mantiene consistente a través de sus distintas manifestaciones. Finalmente, obliga a replantear la noción de percepción judicial, pues lo que el juez observa no es el hecho digital en sí, sino una actualización contingente de este en un entorno técnico determinado.

---

[12] Entiéndase aquí reglas técnicas en un sentido informático (no en el sentido normativo de von Wright).

[13] Mason (2025). El ejemplo fue originalmente formulado por James Christie y reproducido por Mason (2023), pp. 34-35. Se trata de un documento emitido en el contexto del *call for evidence* abierto por el Ministerio de Justicia del Reino Unido (2025) entre enero y abril de 2025 para revisar la presunción del *common law* según la cual los sistemas informáticos operan correctamente, salvo prueba en contrario. El escándalo *Post Office Horizon* (1999-2015), en el que defectos sistémicos del software Horizon llevaron a la persecución penal de más de novecientos subpostmasters británicos, en condenas fundadas en dicha presunción, nos permite ilustrar los problemas que genera la prueba digital cuando no se atiende a su naturaleza y se aborda de manera superficial con las categorías clásicas. Para un análisis detallado, véase Mason y Seng (2021), cap. 5.

Esta característica desestabiliza las categorías espaciales propias del derecho probatorio tradicional y, además, impacta directamente en la forma en que estos hechos se inscriben y persisten en el tiempo. En la medida en que un mismo evento informacional puede desplegarse simultáneamente en múltiples instancias técnicas y permanecer disponible más allá del momento de su ocurrencia, la temporalidad deja de ser un dato externo que se reconstruye retrospectivamente y pasa a convertirse en una dimensión constitutiva del propio hecho digital.

3.5. Temporalidad difusa

A diferencia del hecho analógico, cuya inscripción temporal suele reconstruirse *ex post* mediante testigos, documentos, etc., el hecho digital se produce y conserva dentro de un régimen temporal técnicamente estructurado. No acontece simplemente en el tiempo para luego desaparecer, sino que se constituye a través de arquitecturas temporales específicas que lo sustentan: relojes del sistema, marcas de tiempo, sincronizaciones de red, latencias, ciclos de actualización, registros de eventos y rutinas automáticas de conservación o borrado.

Además, el hecho digital no se inscribe en una temporalidad homogénea y lineal, sino en lo que Rob Kitchin denomina *timescapes* (cronopaisajes) digitales: paisajes temporales complejos, compuestos por múltiples capas, ritmos y escalas de tiempo que coexisten y se superponen dentro de un mismo sistema sociotécnico. En estos entornos, el pasado, el presente y el futuro no siguen una secuencia fija, sino que se superponen y reconfiguran continuamente mediante operaciones de registro, almacenamiento, actualización y reutilización de la información.

Así, un mismo hecho digital puede participar simultáneamente de distintas modalidades temporales. Puede ser, al mismo tiempo, un presente-presente (registro en tiempo real de un evento), un pasado-presente (dato archivado que permanece disponible y operativo), y un presente-futuro (insumo para sistemas de anticipación, modelación o decisión automatizada) [14].

El tiempo digital se puede representar de mejor manera a través de un diagrama de flujo que a través de una simple línea recta con un punto *a quo* y un punto *ad quem*.

Mientras la temporalidad analógica suele ser reconstruida bajo un esquema lineal, la temporalidad digital tiende a organizarse como una temporalidad difusa, estratificada y dependiente de la arquitectura técnica.

Esta transformación tiene consecuencias epistémicas relevantes para el proceso judicial. En el mundo analógico, la pregunta acerca del tiempo suele girar en torno a cuándo ocurrió un hecho. En el mundo digital, la pregunta se desplaza hacia cómo

[14] Para revisar un esquema completo de los ritmos, relaciones y modalidades temporales planteadas por el autor, *cfr*. Kitchin (2023), p. 6.

se produjo, fijó y estabilizó su temporalidad difusa desde un punto de vista técnico. No basta con afirmar que un registro tiene fecha y hora; es necesario examinar el régimen temporal del sistema que genera esas marcas: su grado de sincronización, sus márgenes de desfase, sus políticas de actualización, sus mecanismos de corrección o sobreescritura. La temporalidad del hecho digital es, en consecuencia, una propiedad sistémica, no un atributo aislado.

Desde el punto de vista probatorio, esto implica que la valoración judicial de los hechos digitales no puede apoyarse en una concepción ingenua del tiempo como parámetro neutral y uniforme. El juez ya no se enfrenta a un pasado cerrado que debe reconstruir, sino a un entramado temporal activo, donde los registros del pasado siguen operando en el presente y condicionando escenarios futuros. La prueba digital introduce una temporalidad espesa, técnica y estratificada, que exige criterios específicos de análisis y control.

3.6. Una singularidad entre singularidades: estratificación en un sistema de planos o multiplanar

La estratificación en planos constituye una singularidad especialmente relevante, pues permite comprender cómo las demás propiedades se articulan dentro de una misma arquitectura técnica. Sabemos que el hecho digital no se reduce a lo que «aparece» en una pantalla, ni a la forma en que es presentado ante el usuario. Esa presentación constituye apenas un modo de acceso, esto es, una actualización perceptible de un fenómeno que se sostiene en determinaciones técnicas heterogéneas.

Para evitar el riesgo de abstracción y porque el análisis estratificado se comprende mejor a través de un caso suficientemente complejo, conviene anclar la exposición de esta singularidad en el siguiente ejemplo hipotético paradigmático:

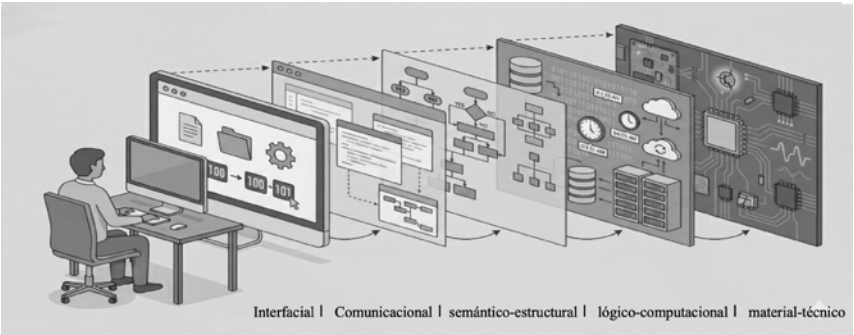
> Un usuario desliza su dedo sobre la pantalla táctil de su dispositivo móvil para marcar «me gusta» en una publicación de una red social. Desde su perspectiva, se trata de un gesto simple, intencional y prácticamente instantáneo. En el dispositivo, el contacto físico con la pantalla genera variaciones en el campo capacitivo del panel táctil y señales eléctricas que son interpretadas por el sistema operativo y por la aplicación. La aplicación traduce esa instrucción conforme a sus reglas internas, identifica la acción como un «me gusta» dirigido a una publicación específica y produce una respuesta inmediata en la interfaz, normalmente la activación visual del ícono correspondiente. La aplicación transmite luego una solicitud técnica estructurada hacia los servidores de la plataforma, identificando al usuario, el objeto digital involucrado y el tipo de acción ejecutada. En el entorno del servidor, esa solicitud es validada y procesada conforme a las reglas internas de la plataforma. Si dichas condiciones se satisfacen, el estado del objeto digital —la publicación— se modifica conforme al modelo de datos del sistema. Luego, el sistema comunica el resultado al dispositivo del usuario y replica progresivamente el nuevo estado hacia otros nodos.

Como se puede apreciar, este ejemplo no describe una secuencia de etapas ni un recorrido lineal del hecho digital que termine fragmentando el proceso en múltiples hechos distintos. Por el contrario, pone de manifiesto que un mismo hecho digital

ocurre de manera concomitante, realizándose bajo distintas formas técnicas que coexisten y se reiteran en diferentes puntos de la arquitectura del sistema. El «me gusta» no corresponde únicamente al gesto físico del usuario ni a la visualización interfacial del ícono, sino a una actualización informacional distribuida, compuesta por operaciones materiales, lógico-computacionales, semántico-estructurales, comunicacionales e interfaciales.

Como veremos a continuación el hecho digital involucra, al menos [15], una dimensión material de soporte, una dimensión lógico-computacional de procesamiento, una dimensión estructural de organización del objeto digital, una dimensión comunicacional vinculada a su distribución y sincronización, y una dimensión interfacial mediante la cual el sistema traduce estados internos a una experiencia accesible para el usuario. Estas dimensiones no operan de forma secuencial ni jerárquica, sino como criterios simultáneos de realización del evento informacional [16].

En la siguiente imagen graficaremos los cinco planos para una mejor comprensión [17]:



(a) Plano material-técnico (infraestructura física)

Corresponde al sustrato físico del hecho digital, es decir, el sustento principal dentro del mundo analógico: dispositivos, sensores, servidores, soportes de almacenamiento, circuitos electrónicos, relojes de sistema y flujos de energía. En este nivel, el hecho digital existe como variación de estado físico, por ejemplo, un transistor que

[15] Algunos autores, como Costa (2021), p. 45, dividen este plano en dos: (a) el *hardware* (servidores, módems, etc.) y (b) la infraestructura (tendidos de cables, satélites, cables submarinos). Sin embargo, preferimos tratar ambas dimensiones dentro de un solo plano, para efectos de simplificar la presentación de este elemento.

[16] La identidad del hecho digital no se agota en ninguno de estos planos, sino en la coherencia estructural que mantiene a través de ellos.

[17] Imagen generada con un sistema de inteligencia artificial generativa (Gemini) con fines exclusivamente ilustrativos. La indicación (*prompt*) dada al sistema requirió una representación esquemática y de estilo sobrio de los planos de ocurrencia del hecho digital, en la forma aquí desarrollada.

pasa de estado de conducción a no conducción, una variación de voltaje registrada por un sensor de una cámara, un micrófono o un Gps.

(b) Plano lógico-computacional (procesamiento y reglas internas)

Consiste en el nivel del *software* ejecutable, *firmware*, algoritmos, modelos de datos y reglas de actualización. El hecho digital surge aquí como una transición de estado lógico, producida conforme a reglas formales: validación de sesiones, ejecución de funciones, asignación de estados, aplicación de filtros automáticos.

Cuando el usuario presiona «publicar», en la capa lógico-computacional el sistema no comunica ni representa aún un mensaje, sino que ejecuta una secuencia de reglas formales: verifica la validez de la sesión del usuario, valida el formato del contenido conforme a parámetros predefinidos, decide mediante funciones internas la creación o no de un registro, aplica filtros automáticos —incluidos mecanismos de moderación, detección de *spam* o intervención algorítmica—, asigna estados lógicos al resultado (v.g., publicado, pendiente o bloqueado) y, finalmente, inscribe ese resultado de acuerdo con el modelo de datos del sistema. En este nivel, el hecho digital surge exclusivamente como una transición de estado lógico, independiente aún de toda significación.

(c) Plano semántico-estructural (objeto digital)

En este plano, el hecho digital se individúa como objeto significativo: publicación, registro, evento, transacción, imagen, video o *log*. Se determina qué cuenta como ese hecho y no otro, conforme a estructuras de datos y convenciones del sistema. Aquí se determina qué cuenta efectivamente como un «me gusta» dentro del universo técnico de la plataforma, distinguiendo una acción válida de una acción fallida, duplicada o inválida, y fijando la identidad del objeto digital resultante.

(d) Plano comunicacional o de red (transmisión y sincronización)

Se trata de un plano en el que el hecho digital se transmite, replica y sincroniza a través de redes técnicas: protocolos de comunicación, latencias, copias distribuidas, respaldos y actualizaciones en distintos nodos. En el ejemplo, la acción de dar «me gusta» no constituye un evento local ni instantáneo, sino un proceso distribuido que se despliega a través de la infraestructura de red del sistema. El gesto del usuario es transmitido desde el dispositivo de origen hacia los servidores de la plataforma —por ejemplo, los servidores de *Meta Platforms, Inc.*— mediante protocolos de comunicación que encapsulan la interacción como una solicitud técnica. En dichos servidores, la solicitud es recibida, validada y procesada conforme a las reglas internas del sistema: se verifica la autenticidad del usuario, se comprueba la coherencia del evento con el estado previo del objeto (por ejemplo, que el «me gusta» no haya sido previamente registrado) y se actualiza el estado lógico correspondiente en las bases de datos del sistema.

Una vez validado y registrado el cambio de estado, el sistema lo propaga hacia múltiples destinos. En primer lugar, comunica el nuevo estado al dispositivo del propio usuario, permitiendo que la interfaz refleje la acción realizada —v.g., mediante la activación del ícono y la actualización del contador—. En segundo lugar, el cambio se replica y sincroniza progresivamente hacia otros nodos del sistema, de modo que el nuevo estado del objeto digital pueda ser accesible para todos aquellos usuarios que, conforme a las reglas de visibilidad de la plataforma, tengan acceso a la publicación.

Este proceso de propagación no es necesariamente simultáneo ni homogéneo. Dependiendo de factores como latencias de red, mecanismos de cacheo, balanceo de carga o replicación asincrónica de bases de datos, distintas instancias del sistema pueden reflejar el cambio en momentos distintos. Así, el «me gusta» puede ser visible de inmediato para algunos usuarios y con retraso para otros, sin que ello implique la existencia de hechos digitales distintos, sino múltiples manifestaciones temporales de un mismo evento informacional distribuido.

(e) Plano interfacial (visualización y acceso humano)

Corresponde a la interfaz gráfica o visual mediante la cual el usuario accede al objeto digital: pantallas, *dashboards*, reproductores, reportes renderizados. El hecho ocurre interfacialmente cuando el ícono de «me gusta» se activa visualmente, el contador se actualiza y el sistema confirma perceptivamente la acción realizada.

El recorrido del ejemplo permite comprender con mayor precisión la naturaleza del hecho digital como evento técnico distribuido. Un único gesto —el deslizamiento del dedo sobre una pantalla táctil— basta para activar, de manera casi instantánea, una pluralidad de operaciones que se realizan simultáneamente en distintos planos, tanto en el dispositivo del usuario como en los sistemas remotos de la plataforma. No se trata de que el hecho «nazca» en el dispositivo y luego «se traslade» al servidor, sino de que el mismo evento se actualiza reiteradamente en ambos extremos de la arquitectura técnica, bajo formas análogas, pero no idénticas.

En el dispositivo del usuario, el hecho ocurre materialmente como variación física, lógicamente como interpretación formal del *input*, semánticamente como tipificación provisional de la acción, e interfacialmente como respuesta perceptible inmediata. En los servidores de la plataforma, ese mismo evento vuelve a ocurrir —no como mera continuación, sino como reconstitución técnica— activando nuevamente operaciones materiales, lógico-computacionales y semántico-estructurales que determinan si la acción cuenta efectivamente como un «me gusta» válido y estable dentro del sistema. El retorno de la información al dispositivo del usuario y su posterior replicación hacia otros nodos no hacen sino reiterar este esquema: en cada punto relevante del sistema, el evento vuelve a actualizarse bajo los mismos planos de ocurrencia.

Este viaje de ida y vuelta muestra que los planos aquí distinguidos no son lugares fijos ni compartimentos estancos, sino formas técnicas de realización del evento que pueden activarse múltiples veces en distintos nodos de la arquitectura digital.

El *truthmaker* digital no reside ni en el gesto inicial, ni en la base de datos, ni en la visualización final, sino en la coherencia dinámica de estas múltiples actualizaciones distribuidas. No es un *truthmaker* plano como el analógico, sino precisamente multiplanar, donde cada plano aporta información sobre su valor epistémico y potencial verificatorio.

4. DESLINDES CON LAS CATEGORÍAS PROBATORIAS TRADICIONALES

Si las propiedades de los hechos digitales se consideran a la luz de las categorías tradicionales del derecho probatorio, es posible advertir ciertas similitudes parciales con nociones ya conocidas, en particular con los hechos institucionales y los hechos técnicos. Sin embargo, una consideración más detenida muestra que el hecho digital no encaja plenamente en ninguna de estas categorías, aun cuando incorpore elementos de ambas, lo que pone de manifiesto una insuficiencia estructural del paradigma probatorio tradicional.

En efecto, el derecho probatorio ha desarrollado históricamente sus categorías fundamentales a partir de una tipología relativamente estable de hechos relevantes: hechos naturales o brutos, hechos técnicos y hechos institucionales. Cada una de estas categorías responde a determinadas condiciones ontológicas y epistémicas que han permitido articular reglas de admisión, producción y valoración de la prueba. El problema que aquí se plantea es si esas categorías tienen realmente la capacidad explicativa o rendimiento suficiente para hacer frente al fenómeno digital, con las múltiples singularidades que se han desarrollado.

El análisis que sigue se propone, en primer lugar, examinar la categoría de hecho técnico y mostrar por qué el hecho digital no puede ser subsumido en ella sin pérdida de su real significado. En segundo lugar, se abordará la noción de hecho institucional, destacando tanto las afinidades como los límites de su aplicación al fenómeno digital. Finalmente, se sostendrá que el hecho digital exige el reconocimiento de una categoría ontológica autónoma, que integre elementos técnicos e institucionales, y que permita fundar una epistemología probatoria adecuada a su naturaleza.

4.1. El hecho técnico y sus límites explicativos

Los hechos técnicos suelen entenderse como aquellos cuya constatación o comprensión requiere el auxilio de conocimientos especializados ajenos a la experiencia ordinaria del juez. Se trata de hechos cuya existencia o características no son directamente accesibles mediante la percepción común, y que deben ser inferidas a partir del funcionamiento de artefactos, instrumentos de medición o sistemas técnicos específicos. La prueba pericial aparece, en este contexto, como el medio idóneo y necesario para su incorporación y valoración en el proceso judicial.

En estos casos, el fenómeno relevante existe con independencia del sistema técnico que lo detecta o mide. La tecnología es un instrumento utilizado para acceder, describir o cuantificar un estado de cosas que preexiste en el mundo, aunque no sea inmediatamente observable. El termómetro no crea la temperatura ni el radar crea la velocidad; solo son instrumentos que miden magnitudes. El hecho técnico presupone un sustrato empírico externo cuya existencia no depende del dispositivo que lo registra.

En este sentido, siguiendo la teoría de Simondon, el objeto técnico está sometido a una génesis y su individualidad se modifica en el transcurso de ese proceso[18]; su especificidad no puede definirse exclusivamente por el uso práctico al que responde, pues ninguna estructura fija corresponde de manera definitiva a un uso determinado. El hecho técnico, lejos de suponer un objeto aislado o un resultado puntual, apunta al acontecer de un sistema en acto, desplegado dentro de un medio asociado que condiciona su funcionamiento.

El hecho digital comparte ciertos rasgos con esta categoría. Su conocimiento exige competencias técnicas y su análisis suele requerir peritajes especializados. Ello se debe a que buena parte de sus determinaciones no son perceptibles directamente y en su producción intervienen tecnologías heterogéneas. Sin embargo, estas similitudes no autorizan su reducción al hecho técnico clásico. A diferencia de este, el hecho digital no presupone necesariamente un fenómeno preexistente que el sistema técnico se limite a medir o representar. En numerosos casos —especialmente en los hechos nativamente digitales— el evento no existe fuera de la arquitectura técnica que lo produce, registra y estabiliza.

Cuando un usuario publica contenido en una red social o cuando un sistema genera una imagen sintética, el hecho relevante no es un estado del mundo analógico que luego se describe técnicamente. El hecho consiste en la transición de estado del sistema informático. El sistema, más que detectar el hecho, lo constituye.

Esta diferencia es decisiva para el derecho probatorio. Si el hecho digital pudiera reducirse a un hecho técnico en sentido clásico, bastaría con una adecuada regulación de la prueba pericial para permitir al juez acceder a él como a cualquier fenómeno físico o químico. Sin embargo, en el ámbito digital lo que está en juego no es sólo el correcto funcionamiento de un dispositivo, sino la fiabilidad global de un sistema que produce, valida y presenta hechos como tales. Por estas razones, aunque el hecho digital incorpora una dimensión técnica innegable, no puede ser reducido a la categoría de hecho técnico sin desconocer su especificidad ontológica.

[18] Simondon (2007), p. 41.

4.2. El hecho institucional y sus límites explicativos

Un segundo acercamiento al problema consiste en comparar el hecho digital con la categoría de hecho institucional. En la teoría social desarrollada por John Searle[19], los hechos institucionales se caracterizan por su dependencia de reglas constitutivas, cuya estructura puede formularse como «X cuenta como Y en contexto C». Contraer matrimonio, celebrar un contrato o asumir un cargo público no son hechos naturales, sino que existen en virtud de reglas que los instituyen y que son aceptadas colectivamente dentro de un determinado marco normativo.

El hecho digital presenta afinidades con esta categoría. Al igual que los hechos institucionales, su existencia depende de reglas que determinan qué cuenta como evento válido dentro de un sistema determinado. El inicio de sesión de un usuario, una publicación, un «me gusta» o una transferencia bancaria sólo existen como tales porque un conjunto de reglas define sus condiciones de producción, reconocimiento y efectos. El hecho digital supone transiciones normativamente estructuradas.

Nuevamente, el parecido es solo parcial. A diferencia de los hechos institucionales, las reglas que instituyen los hechos digitales son primariamente técnicas y algorítmicas (en último término, lógicas)[20]. Se trata de reglas inscritas en código, diseñadas por actores privados, opacas para el juez y no directamente gobernadas por el ordenamiento jurídico, aun cuando produzcan efectos jurídicamente relevantes. Mientras los hechos institucionales tradicionales se estabilizan mediante procedimientos normativos relativamente transparentes —actos formales, registros públicos, rituales jurídicos—, los hechos digitales se estabilizan a través de procesos técnicos cuya lógica interna no coincide con las categorías jurídicas de validez, nulidad o inexistencia. El hecho digital no «vale» porque haya sido jurídicamente instituido, sino porque el sistema técnico lo reconoce como tal y lo hace persistir bajo determinadas condiciones de coherencia interna.

Por ello, aunque el hecho digital incorpora una dimensión institucional, en cuanto depende de reglas similares a las constitutivas, no puede ser identificado sin más con un hecho institucional en sentido clásico, en la medida que su institucionalidad es técnica y no jurídica.

---

[19] Searle (2010), p. 90 y Searle (1969) [2011], pp. 34-35. En este mismo sentido, González Lagier (2024), p. 26, define los hechos institucionales como «aquellos descritos o interpretados desde el punto de vista de un sistema de normas».

[20] De hecho, se ha definido el algoritmo como «un conjunto finito de reglas no-ambiguas que se pueden aplicar sistemáticamente a un objeto o a un conjunto de objetos para transformarlos de maneras definidas y circunscritas» Bermúdez (2020), p. 23. En otras palabras, son reglas en un sentido lógico-formal.

4.3. El hecho digital como categoría ontológica autónoma

Las consideraciones precedentes permiten extraer otra conclusión central. El hecho digital no es una mera variante del hecho técnico ni una simple extensión del hecho institucional. Se trata de una categoría ontológica autónoma, que integra elementos técnicos e institucionales sin agotarse en ninguno de ellos. Su especificidad se encuentra en el hecho de que constituye un evento informacional producido, registrado y estabilizado dentro de arquitecturas técnicas que definen, desde su origen, las condiciones de su existencia, accesibilidad y persistencia.

El hecho técnico presupone un fenómeno externo que el dispositivo permite conocer; el hecho institucional presupone una regla social o jurídica que confiere estatus; el hecho digital, en cambio, presupone una arquitectura técnica que produce, inscribe y estabiliza una transición de estado. Su singularidad consiste en que la regla no se limita a calificar un hecho previo ni a medir un fenómeno externo, sino que participa en la constitución misma del acontecimiento informacional.

5. DE LA EPISTEMOLOGÍA DE LA PRESENCIA A UNA EPISTEMOLOGÍA DE LA PRODUCCIÓN TÉCNICA

El giro epistemológico que proponemos se apoya en la distinción clásica de Russell entre *knowledge by acquaintance* y *knowledge by description*[21]. En su formulación, el primero no depende de inferencias ni de descripciones, sino de una relación epistémica originaria en que el objeto se da o «se presenta» ante el sujeto:

> «decir que S tiene *acquaintance* con O es esencialmente lo mismo que decir que O es presentado a S».

Esta equivalencia, que da cuenta de un tránsito bidireccional entre S y O, no describe un estado psicológico del sujeto, sino una estructura de acceso en la cual el contenido del conocimiento queda anclado en el modo de comparecencia del objeto. Lo decisivo no es qué cree el sujeto, sino qué se le da y bajo qué forma. La epistemología de la presencia presupone que el objeto del conocimiento puede comparecer como constituyente del juicio, ya sea de manera inmediata o mediante una mediación todavía recorrible. En el proceso judicial clásico, esta estructura se reproduce: el juez oye al testigo, observa el objeto, examina el documento. Incluso cuando existe mediación técnica, subsiste la idea de que el hecho es, en último término, algo que ocurrió independientemente del medio que lo representa.

El conocimiento por descripción, en cambio, aparece precisamente allí donde el objeto no se presenta. En estos casos, el sujeto solo puede referirlo de manera mediata o inferencial, a través de proposiciones, signos, reportes o cadenas explicativas. La

[21] Russell (1911), pp. 112-113.

justificación ya no reposa entonces en la presencia del objeto, sino en factores sustitutivos: la credibilidad del informante, la estabilidad del soporte, la corrección del procedimiento y la consistencia del sistema que permite la referencia. Russell introduce esta distinción para mostrar que una parte sustantiva de nuestro conocimiento —incluido el conocimiento de objetos físicos y de otras mentes— no puede fundarse en la percepción directa, sino que depende de estructuras descriptivas que sustituyen la presentación por mecanismos de referencia.

Pues bien, la práctica judicial pasa por alto que los hechos digitales se producen mediante descripciones técnicas; no se encuentra en el mundo analógico esperando a que el sujeto las perciba. La visualización en pantalla se asume como si mostrara el hecho, cuando en realidad lo que comparece ante el sujeto es el reporte de un sistema sobre un estado lógico, traducido a una interfaz inteligible para humanos.

En el entorno digital el *knowledge by description* se ve llevado a un extremo inédito: no solo carecemos de *acquaintance* con relación al objeto, sino que el propio objeto carece de una existencia independiente del sistema técnico que lo produce. Esto nos lleva a dos axiomas digitales que podemos formular como: *nullum obiectum sine systemate* (el objeto digital es inescindible del sistema que lo produce) y *nullum systema sine descriptionibus* (el sistema es inconcebible sin un conjunto coordinado de descripciones).

El impacto de la epistemología de la producción técnica se extiende incluso al ámbito lógico-lingüístico, pues las proposiciones acerca de un hecho digital deben ser reformuladas en términos análogos a los ya establecidos para las descripciones definidas. Así, deben revestir más bien una sintaxis del tipo:

> "que existe exactamente un x tal que x satisface las condiciones técnicas $C_1 \ldots C_n$, y que ese x es tal-y-tal".

5.1. Tres regímenes de comparecencia del hecho

Dicho lo anterior, proponemos distinguir tres regímenes ontológico-epistémicos de comparecencia del hecho, relevantes para la teoría de la prueba.

En primer lugar, (a) un régimen de *presentación directa*, en el cual el hecho existe como evento inmediato y comparece sin mediación técnica, de modo que la epistemología puede todavía orientarse por un modelo de presencia.

En segundo lugar, (b) un régimen de *representación técnica*, en el que el hecho preexiste, pero comparece mediante una inscripción causal material —huella, medición, registro científico-analógico—, de manera que el conocimiento ya no consiste en la presencia del hecho, sino en su reconstrucción a partir de una representación causal relativamente recorrible. En ambos casos, aun cuando exista mediación, el objeto mantiene una referencia externa que permite, al menos en principio, un recorrido epistémico desde la representación hacia el evento.

El (c) tercer régimen, propio del entorno digital, es el de *producción técnica*. Aquí el hecho no se presenta ni se representa, sino que se genera dentro de una arquitectura técnica mediante reglas de captura, procesamiento, validación y visualización, y comparece ante el sujeto (juez, perito, etc.) como un resultado renderizado por el sistema. En este régimen, la relación epistémica se invierte: ya no es el sujeto quien avanza hacia el objeto para reconstruirlo, sino el objeto —producido técnicamente— el que emerge hacia el sujeto como un hecho consumado. La comparecencia del hecho no depende de la percepción ni de la reconstrucción, sino de la forma en que el sistema decide hacerlo aparecer. Lo que llega al sujeto es un *output* ya configurado. La interfaz emula a la epistemología de la presencia y define qué estados del sistema se hacen visibles, bajo qué forma se organizan, qué relaciones quedan accesibles y cuáles permanecen fuera del campo de interrogación epistémica.

La interfaz opera como una instancia constitutiva del acceso al hecho digital. La confianza espontánea en lo visible —heredada del paradigma analógico— se convierte aquí en una fuente de error, pues induce a tratar como dado aquello que ha sido previamente seleccionado, filtrado y estabilizado por el sistema.

5.2. De la correspondencia a la producción

De lo anteriormente expuesto se sigue una consecuencia central para la epistemología probatoria. La prueba digital no puede fundarse en la epistemología de la presencia, porque los hechos que ingresan al proceso judicial ya no se dan como eventos preexistentes que el juez deba descubrir o reconstruir. El conocimiento no se orienta solamente a establecer correspondencias con un pasado empírico, sino también a evaluar la fiabilidad de los procesos técnicos que producen ese estado informacional. La distinción entre *knowledge by acquaintance* y *knowledge by description* es crucial porque el derecho probatorio suele tratar la visualización en pantalla como si fuera *acquaintance* («es verdad porque lo vi»), cuando en realidad es *description* («es verdad porque lo inferí»).

La legitimidad de la decisión no se apoya en el acceso directo al hecho ni en la mera autoridad técnica del experto, sino en la capacidad institucional del juez para controlar, auditar y estabilizar críticamente la generación y renderización del hecho digital, lo que exige una adaptación de la epistemología probatoria, de las reglas de valoración de la prueba y, en último término, de la racionalidad judicial misma.

Ahora bien, frente a la insuficiencia de la correspondencia, la limitación estructural de las percepciones, y a la indefectible opacidad algorítmica, ¿cómo puede el derecho reconstruir una epistemología operativa? La respuesta no puede ser un retroceso hacia la *verifobia*[22] (negar valor a la prueba digital), sino más bien un racionalismo

[22] Taruffo (2014), p. 32, señala que la verifobia se refiere a «todas aquellas actitudes que, de manera más o menos abierta y consciente, son contrarias a cualquier discurso que reconozca significado y valor a

escéptico que, mediante un ejercicio abductivo de contracción, pueda estrechar el margen de incertidumbre[23].

6. CONSECUENCIAS PARA LA TEORÍA DE LA PRUEBA

A continuación, veremos algunas consecuencias relevantes para la teoría de la prueba que ilustran sobre el potencial de lo desarrollado hasta ahora.

6.1. El colapso del paradigma del cotejo

El derecho probatorio parece haber perdido el foco al considerar la prueba digital (únicamente) como documento. Además, la determinación de la autenticidad se ha entendido en el sentido de cotejar una copia con su original o la conservación material de un soporte estable cuya integridad puede verificarse mediante control de cadena de custodia o examen pericial[24].

Sin embargo, este esquema presupone tres condiciones: la existencia de un objeto unitario, su localización espacial relativamente determinada y su estabilidad material a lo largo del tiempo. Si el hecho digital es, en cambio, una transición de estado distribuida, replicada y estratificada, esas presuposiciones dejan de ser estructuralmente sostenibles.

La noción de autenticidad depende, en realidad, de la fiabilidad del proceso técnico que produjo y estabilizó el evento. En este contexto, el paradigma del cotejo no se muestra simplemente insuficiente, sino epistemológicamente desajustado.

6.2. La crisis del modelo perceptivo de valoración

La valoración judicial de la prueba se ha apoyado históricamente en un modelo que hemos denominado epistemología de la presencia, con objetos que comparecen directamente ante el sujeto. En el entorno digital, esta estructura se altera de manera decisiva. Aquí el juez no percibe el hecho en sí, sino una actualización renderizada

---

la verdad y, por lo tanto, negar de vez en cuando la existencia, la posibilidad, la oportunidad de buscarla y la posibilidad de descubrirla».

[23] Esta forma de razonar, que denominaremos *ratio contractiva*, será desarrollada en un trabajo posterior.

[24] Así se logra apreciar, por ejemplo, en Meza (2024, p.2) y Abdullah, Maqsood y Nadeem (2025), p.795, quienes abordan la problemática de la prueba digital desde las categorías clásicas de autenticidad, integridad, cadena de custodia y confiabilidad de las fuentes. Este enfoque —frecuente en la literatura reciente— presupone que la evidencia digital constituye una extensión tecnológica de los medios probatorios. tradicionales y que, en consecuencia, sus desafíos pueden resolverse mediante una intensificación del razonamiento probatorio dentro de esas mismas categorías

por el sistema. La interfaz no constituye una ventana transparente hacia un evento, sino una instancia de producción, selección y organización de información.

El supuesto de que la visualización equivale a conocimiento directo se debilita. No se trata de una ilusión perceptiva en sentido clásico, sino más bien de una mediación estructural. El juez solo accede a una manifestación parcial cuya inteligibilidad depende de la arquitectura que la hace aparecer.

6.3. La insuficiencia de la prueba pericial

Podría sostenerse que las dificultades anteriores se superan mediante una ampliación o sofisticación de la prueba pericial. En efecto, la pericia es útil para esclarecer aspectos técnicos del sistema: su funcionamiento, reglas de registro, coherencia entre planos e integridad formal de los objetos digitales examinados. Sin embargo, esa utilidad no agota el problema probatorio. Cuando el hecho consiste en una transición de estado producida dentro de una arquitectura informacional, el informe técnico puede explicar cómo opera el sistema, pero no determina por sí solo si el evento satisface las exigencias justificativas del proceso[25].

Además, no es posible exigir prueba pericial en todo caso en que intervenga un hecho digital. Su necesidad dependerá del contexto procesal y del nivel de controversia técnica: a mayores exigencias contextuales, mayor será la necesidad de esclarecimiento. Con todo, incluso entonces, la pericia no resuelve íntegramente el problema, pues la valoración excede el mero análisis técnico.

6.4. La verificación arquitectónica a partir del ***truthmaker*** digital

Los hechos analógicos, que constituyen el clásico *truthmaker* del mundo físico, no tienen una estructura técnica, no nacen inscriptos, son espacialmente ubicables y temporalmente lineales, son alterables solo en la medida que las leyes de la naturaleza lo permitan y no son esencialmente estratificados.

Si el *truthmaker* es aquello en virtud de lo cual una proposición puede ser verdadera o falsa, la transformación de su estructura no puede ser indiferente para la teoría de la prueba. Cuando el verificador es un hecho digital, la evaluación probatoria no puede limitarse a constatar la apariencia del objeto presentado, sino que debe reconstruir las condiciones arquitectónicas que permiten vincular ese objeto con la transición de estado que se afirma como ocurrida.

---

[25] En la doctrina anglosajona, Roberts y Zuckerman (2010), cap. 7, han sostenido una tesis convergente al subrayar que la admisibilidad y valoración de la prueba digital plantean problemas epistémicos distintos a los de la prueba pericial científica tradicional, precisamente porque la fiabilidad del registro depende de condiciones arquitectónicas que el perito puede describir pero no necesariamente justificar en términos jurídicamente relevantes.

Una verificación arquitectónica exige, por tanto, identificar qué transición de estado se alega, qué objeto digital permite acceder a ella, qué arquitectura técnica la produjo, qué planos del sistema la respaldan y qué riesgos de alteración, sincronización, renderización o pérdida de trazabilidad afectan la inferencia probatoria. La intensidad de esta reconstrucción no será siempre la misma: dependerá del estándar probatorio aplicable, del grado de controversia técnica y de la relevancia que el hecho digital tenga para la decisión.

De este modo, el hecho digital no solo introduce nuevos objetos de prueba, sino una forma distinta de verificar. Su fiabilidad va más allá de la estabilidad del archivo, de la existencia de un *hash* o de la opinión del perito; depende más bien de la coherencia entre los distintos planos, registros y condiciones técnicas que permiten estabilizar el evento verificador de la proposición objeto de prueba.

7. CONCLUSIONES

El diagnóstico desarrollado a lo largo de este trabajo puede sintetizarse en una tesis central: el paradigma probatorio clásico presupone una teoría unitaria de los hechos que no distingue adecuadamente entre el modo de existencia del hecho analógico y el del hecho digital. Esta presuposición conduce a aplicar a los hechos digitales las mismas bases ontológicas y epistemológicas diseñadas para eventos materiales, localizados y pasados.

Sin embargo, hemos mostrado que los hechos digitales presentan seis singularidades estructurales específicas —estructurabilidad técnica, registrabilidad originaria, ubicuidad, temporalidad difusa, alterabilidad y estratificación multiplanar— que los diferencian de manera decisiva de los hechos analógicos. Estas singularidades no son meras variaciones superficiales, sino rasgos constitutivos que afectan la forma en que el hecho existe y, por consiguiente, la forma en que puede ser conocido y justificado. Además, las consecuencias no son marginales. El paradigma del cotejo documental, el modelo perceptivo de valoración, la reducción del problema a la prueba pericial y la verificación arquitectónica a partir del hecho digital muestran límites estructurales frente a la epistemología del hecho digital.

Este trabajo, por tanto, no propone simplemente añadir nuevas reservas técnicas a las categorías probatorias existentes. Su tesis es más radical: cuando un *truthbearer* refiere a un hecho digital, el verificador pertenece a un régimen ontológico distinto del presupuesto por la teoría probatoria clásica. De ello se sigue que la valoración probatoria no puede descansar únicamente en la inspección del objeto visible, en el cotejo documental o en la autoridad del perito, sino en la reconstrucción crítica de las condiciones técnicas que produjeron, registraron y estabilizaron el acontecimiento informacional relevante.

Todo lo anterior representa el punto de partida de una transformación profunda en la racionalidad probatoria y en la manera en que el derecho conoce, estabiliza y decide sobre los hechos digitales.

BIBLIOGRAFÍA

Abel Lluch, Xavier y Picó i Junoy, Joan (2011). La prueba electrónica. Estudios prácticos sobre los medios de prueba. Barcelona: Bosch Formación - Esade.

Amaya, Amalia y Páez, Andrés (2026), *El legado de Larry Laudan y el futuro de la epistemología jurídica*, *Quaestio facti*, núm. 10.

Abdullah, Hafiz Omer, Maqsood, Mudassir y Nadeem, Ahmad (2025). *Digital evidence in criminal proceedings: Legal standards, chain of custody, and evidentiary reliability in the digital era.* Research Journal for Social Affairs, 3.

Acosta-León, Carlos (2023). La prueba documental de fuentes informáticas basadas en documentos digitales con firma electrónica certificada: Un análisis desde la perspectiva de las tecnologías de la información y el principio de libertad de prueba en el proceso penal venezolano. *Revista Chilena de Derecho y Tecnología* (12), pp. 1-40.

Bermúdez, Jorge Luis (2020). *Cognitive Science. An introduction to the Science of the Mind* (Third Edition). Cambridge: Cambridge University Press.

Brighi, Raffaella y Ferrazzano, Michele (2021). Digital Forensics: Best Practices and Perspective. En *Digital Forensics Evidence. Towards Common European Standards in Antifraud Administrative and Criminal Investigations*. (Ed. Caianiello, M. y Camon, A., pp.13-48) Milano: Wolters Kluwer – Cedam.

Cardona, Juan David (2020). *La valoración de la prueba o evidencia digital en los procesos judiciales.* Bogotá: Grupo Editorial Ibañez e Instituto Iberoamericano de derecho digital y de la ciberseguridad.

Costa, Flavia (2021). *Tecnoceno. Algoritmos, biohackers y nuevas formas de vida.* Buenos Aires: Taurus.

Estrada, Eduardo (2025). Ontología y prueba de los hechos digitales: una metodología para guiar la decisión judicial en contextos de opacidad epistémica. *Revista Uruguaya de Derecho Procesal* (2).

Galvis, Ángel y Bustamante, Mónica (2019). La no equivalencia funcional entre la prueba electrónica y la prueba documental: Una lectura desde la regulación procesal colombiana. *Revista Ius et Praxis* (2).

González Lagier, Daniel (2024). Prueba, hechos y verdad. En J. Ferrer, *Manual de Razonamiento Probatorio.* Marcial Pons.

Hui, Yuk (2012). What is a Digital Object? *Metaphilosophy (43)*.

Hui, Yuk (2016). *On the existence of Digital Objects.* Minneapolis: University of Minnesota Press.

Ingvarsson, Jonas (2021). *Towards a Digital Epistemology. Aesthetics and Modes of Thought in Early Modernity and the Present Age.* Cham: Palgrave Macmillan.

Kitchin, Rob (2023). *Digital Timescapes. Technology, Temporality and Society.* Cambridge: Polity Press.

Latour, Bruno (1987). Science in Action. Cambridge, Ma: Harvard University Press. Meza, Matías (2024). Elementos probatorios en entornos digitales. La tensión entre la escasez y la sobreabundancia de información en la etapa de investigación. *Repositorio del Master en Razonamiento Probatorio de la Universidad de Girona*.

Mason, Stephen y Seng, Daniel (2021). *Electronic Evidence and Electronic Signatures* (5° ed.). University of London Press.

Mason, Stephen (2023). Evidence from computers: the unreliable legal presumption that, without more, it can be relied upon. *The Barrister*, 95, 34-35.

Mason, Stephen (2025). *The use of evidence generated by software in criminal proceedings.* Submission to the Ministry of Justice Call for Evidence, 14 de abril de 2025. https://law.nus.edu.sg/trail/wp-content/uploads/sites/9/2025/08/The-use-of-evidence-generated-by-software-in-criminal-proceedings-Mason-for-Trail.pdf

Ministry of Justice (2025*). Use of evidence generated by software in criminal proceedings: Call for evidence.* Gov.Uk. https://www.gov.uk/government/calls-for-evidence/use-of-evidence-generated-by-software-in-criminal-proceedings/use-of-evidence-generated-by-software-in-criminal-proceedings-call-for-evidence

Roberts, Paul y Zuckerman, Adrian (2010). Criminal Evidence (2° ed.). Oxford University Press.

Russell, Bertrand (1911). Knowledge by acquaintance and knowledge by description. *Proceedings of the Aristotelian Society, New Series* (11), pp. 108–128.

Santos, Agustina (2025). Hacia una regulación autónoma de la prueba digital. *Revista de Derecho* (N° 24).

Searle, John R. (2010). *Making the Social World*. Oxford University Press.

Searle, John R. (2011). *Speech Acts. An Essay in the Philosophy of Language*. Cambridge University Press.

Simondon, Gilbert (2007). *El modo de existencia de los objetos técnicos*. Buenos Aires: Prometeo.

Taruffo, Michele (2014). De La verità nel processo. En *Derecho procesal: dilemas sobre la verdad en el proceso judicial* (pp.31-48). Editorial Universidad Pontificia Bolivariana, 2014.

Quaestio facti. Revista Internacional sobre Razonamiento Probatorio
Quaestio facti. International Journal on Evidential Legal Reasoning
Sección: Ensayos
2026 I 11 pp. 89-121
Madrid, 2026
DOI: 10.33115/udg_bib/qf.i11.23272
Marcial Pons Ediciones Jurídicas y Sociales
© Piero Mattei-Gentili
ISSN: 2604-6202
Recibido 22/4/26 | Aceptado 22/05/26 | Publicado *online*: 04/06/2026
Editado bajo licencia Reconocimiento 4.0 Internacional de Creative Commons

PROVING CUSTOMARY NORMS: ONTOLOGY AND EVIDENCE IN INTERNATIONAL LAW

Piero Mattei-Gentili*
Instituto de Investigaciones Jurídicas,
Universidad Nacional Autónoma de México.
pieromg@unam.mx
https://orcid.org/0000-0002-5289-1664

ABSTRACT: This article addresses the epistemological and conceptual challenges involved in proving customary norms in international law. It argues that confusion surrounding *usus* and *opinio juris* stems from a failure to distinguish between the ontology of customary norms and the determination of their existence. Drawing on legal epistemology, social norm theory, and analytical legal philosophy, the article conceptualizes customary norms as products of socially constituted structures grounded in shared normative expectations. Proving *opinio juris* does not require showing that States regard their conduct as legally mandatory, but rather that second-order normative expectations are manifested through patterns of practice. The article defends a dual inferential model — inductive reasoning to establish *usus*, complemented by abductive reasoning to infer *opinio juris*— and proposes a narrative coherence test as a qualitative supplement to the preponderance standard.

KEYWORDS: customary international law, *opinio juris*, legal epistemology, social norms, standards of proof.

SUMARIO: 1. INTRODUCTION.— 2. EPISTEMOLOGICAL AND ONTOLOGICAL FOUNDATIONS.— 3. THE STRUCTURE OF CUSTOMARY NORMS: 3.1. The Conventional-

* I want to thank the diligent and sometimes demanding reviewers of this paper. Their feedback was challenging at times, pushing me to be humble and reevaluate my positions, while also encouraging me to better defend my ideas where needed. As a result, the final article is substantially improved.

ist Framework of Expectations; 3.2. Normative Expectations; 3.3. Normative Attitudes Without Moral Endorsement.— 4. Customary *law*. A theoretical subject.— 5. CONCEPTUAL CLARIFICATION FOR INTERNATIONAL LAW DOCTRINE: 5.1. The Relevance of the *Opinio Juris;* 5.2. The Issue of Inadequate Justifications.— 6. THE EVIDENTIARY METHOD AND APPLICATION: 6.1. Inferring Mental States from Repeated Actions; 6.2. Narrative Interpretative and Approach; 6.3. Towards a Standard of Proof.— 7. CONCLUSION.— BIBLIOGRAPHY

1. INTRODUCTION

Custom, as a source of law, is an intriguing subject for legal theory and doctrine because it forces us to reconsider how we approach legal issues by posing a distinctive epistemic problem. In contrast to statutory sources, the main difference is that, in the face of a controversy, one must first prove that a customary legal norm governs the case; a norm that does not originate in an identifiable authoritative act of positing. That is, while with statutory law the epistemic task of proving is confined to the *quaestio facti* (the determination of the factual matters in controversy), customary law extends the evidentiary inquiry to the *quaestio juris* (the concerns involved in determining the applicable law).

The epistemic burden of establishing the *quaestio juris* underscores the practical necessity for legal theory to help determine whether these norms are part of the law. Unlike statutes, customary norms do not originate from an authoritative act of enactment. For legal positivism, which defines law in terms of social facts, this raises the question: what social facts constitute customary norms, and how can legal institutions identify them? Legal doctrine has developed the «two-element formula» as an answer: for there to be a customary legal norm, one requires both general practice (*usus*) and acceptance as law (*opinio juris*). This provides guidance but leaves unresolved issues, particularly regarding adjudication. First, there is the ontological question about the nature of customary norms: How are we understand the elements of the formula? The structure of proof depends on the structure of what is to be proven. Once we have an answer, we can consider which epistemological method is adequate for proving the existence of customary norms and what makes them part of the law.

The fact that custom is a significant source of international law makes it a compelling case for study, since identifying customary legal norms remains consistently difficult and highlights the conceptual and epistemological challenges that legal theory should address. The latter reflects in the fact that the International Court of Justice and other tribunals apply the two-element formula inconsistently, sometimes emphasizing State practice, sometimes *opinio juris*, without a clear notion of how these elements interact or what evidence suffices to establish either. Scholarship in international law has addressed these difficulties primarily through internal doctrinal refinements, such as distinguishing «traditional» from «modern» custom, proposing «sliding-scale» approaches in which strong evidence of one element compensates for weak evidence of the other, and debating whether *opinio juris* requires belief that

conduct is legally obligatory or merely normatively expected. However, these debates are conceptually obscure and could benefit from incorporating insights from legal theory and conceptual analysis[1].

This article contends that the confusion in methodology regarding customary international law stems from insufficient attention to the connection between ontology and epistemology. Scholars and courts conflate questions about what constitutes a customary norm with questions about how to prove its existence. This reasoning fuses the evidentiary outcome with a conclusion about the ontological requirements for the existence of norms. Courts' methodological choices — what evidence they emphasize — do not necessarily reveal what norms *are*. The result is conceptual inconsistency: courts apply the «two-element formula» erratically, and, consequently, the doctrine offers erratic ontological accounts that paper over rather than resolve the underlying confusion. Drawing on work in social ontology, norm theory, and legal epistemology, I develop a framework that reconceptualizes the relationship between *usus* and *opinio juris* and proposes narrative reconstruction as the appropriate method for proving customary norms. Customary norms are conceptual objects: they exist in virtue of participants' normative expectations. This ontological clarification has epistemological implications. Proving customary norms cannot be a matter of simple empirical observation or inductive generalization from instances of State practice. Rather, it also requires abductive inference. That is, reasoning from recurrent observable phenomena to the best explanation available for the peculiar manifestation of certain reactions (normative expectations) within that structure of concurrent State practices[2].

The article presents three contributions. First, it provides a conceptually rigorous account of the relationship between *usus* and *opinio juris*. I argue that these are not variables requiring separate proof but rather parts of a single inferential procedure: practice serves as evidence from which normative expectations can be inferred. This can be regarded as a solution for the so-called «chronological paradox» (how can

[1] Situated within international legal doctrine, D'Amato observes that although some international law scholars address legal theory, their engagement frequently lacks sufficient depth. D'Amato, 1971.

[2] 2 Charles S. Peirce explained that abductive reasoning involves a form of pragmatic reasoning that infers causes from effects. In its more contemporary form, it is also called *inference to the best explanation,* and, as Lipton proposes, it can be seen as a form of reasoning in which the phenomenon to be explained provides essential input for believing that a certain explanation is correct (Lipton, 2000). In our case, the very form of States' conduct supports the idea that normative expectations best explain the phenomenon. It includes normative cues, like justifying actions and reactions to deviations from routines. Hence, it is valid to wonder why States justify how they act, why they react to others' divergent actions? We seek the underlying cause as the best explanation possible. Observing these reactions reflects reciprocal expectations about appropriate actions and about what must be done; we can infer that a normative framework — specifically, a customary norm among these States — is at play. It is also worth noting, as Jones does, that abductive inference is commonly employed, particularly in the social sciences, to interpret the existence of social beliefs (such as norms) and to assume certain rationality among agents (Jones, 2011, pp. 1701-1702).

practice manifest *opinio juris* if it is necessary for practice to count as legally relevant?) and explains why doctrinal distinctions between «traditional» and «modern» custom, or proposals for «sliding-scale» methodologies, rest on conceptual confusion. Second, the framework clarifies what courts should be looking for when they seek *opinio juris*: not beliefs that conduct is already law, but normative expectations. Third, I propose a methodological approach to proving customary norms: narrative reconstruction, which infers normative expectations by coherently integrating diverse evidence, including State practice, official statements, patterns of approval and criticism, and responses to perceived violations.

The article proceeds as follows. Part II establishes the epistemological foundations, explaining why proving social norms requires special attention to their ontological status, dependent on participants' normative attitudes. Part III draws on social norm theory to clarify the structure of customary norms, distinguishing normative expectations from mere empirical regularities. Part IV addresses the relationship between social norms and legal norms, explaining how custom becomes law without requiring a separate «act of positing.» Parts V and VI apply these theoretical insights to resolve persistent confusions in international law doctrine and develop concrete guidance for evidentiary reasoning in courts. Part VII concludes by reflecting on implications for legal practice and legal theory.

2. EPISTEMOLOGICAL AND ONTOLOGICAL FOUNDATIONS

Knowledge requires more than true belief; it requires justified belief grounded in intersubjective reasons[3]. The reasons justifying our claims must be accessible to any rational agent and subject to critical evaluation through shared criteria of validity. If making a claim entails commitment to its truth, then having a firm internal conviction does not constitute a valid epistemological reason for qualifying such a belief as knowledge. Intersubjective criteria assess the truth-validity of any claim. Claims about customary norms must be measured against external reality, presupposing a correspondence theory of truth (Goldman, 1999, pp. 59-65; Ferrer Beltrán, 2005, pp.73-75; Haack, 2014, p. 302). Yet we often cannot observe social phenomena directly, and evidence-based reasoning is inherently probabilistic. No set of evidence, however rich and reliable, allows us to reach certainty about a phenomenon's occurrence. Therefore, proving a customary norm exists means showing it probably exists given the available evidence (Ferrer Beltrán, 2007, pp. 92-93).

The epistemological challenge is compounded by the ontology of customary norms. Unlike material objects or directly observable events, such as brute facts,

[3] I will employ the «true justified belief» model of knowledge as a working assumption. I am aware that, since Gettier, this model has been shown to be imperfect, and that philosophers such as Williamson have argued against it, it remains the standard philosophical blueprint for epistemological analysis (Williamson, 2002).

customary norms are not mind-independent facts. They are conceptual objects[4]. If they exist, they do so in virtue of participants' normative attitudes within a recurrent practice — specifically, expectations about how participants ought to behave. This ontological feature, from an epistemological perspective, poses a methodological requirement to determine the presence of a shared mental state (normative expectations), not merely behavioral regularities. Hence, one must infer its existence from evidence of participants' attitudes as manifested in their conduct (Mattei-Gentili, 2026, pp. 53-54). The absence of epistemically valid methods for making such inferences increases the risk of erroneous determinations — mistaking mere regularities for norms or failing to recognize existing norms[5].

As Susan Haack observes, institutions exist in virtue of what people collectively do and believe. They are socially dependent, yet real enough to ground objective statements about them (2014, pp. 309-310). Norms, as social institutions, share this ontological status: they are neither purely subjective (dependent on any individual's beliefs) nor mind-independent (existing regardless of participants' attitudes). Rather, they exist when and because participants share certain expectations. Understanding this ontological structure is essential for determining which evidence can justify claims about the existence of customary norms. Conceptual analysis clarifies what social facts constitute them — the conditions under which such norms exist[6]. Legal theory then addresses the conditions under which social norms become law. These inquiries are necessary for a complete account of customary law because they address different questions: ontology (what are norms?), epistemology (how do we know them?), and legal theory (when are they law?).

3. THE STRUCTURE OF CUSTOMARY NORMS

Common legal doctrine tends to rely primarily on written sources when addressing legal issues. Traditionally, analytical legal theory has assumed a similar scope, treating legal norms as part of a prescriptive discourse (Celano, 2010: 174). From this perspective, international law's reliance on custom seems odd, even raising doubts about whether it is law at all (Hart, 1994, p. 227).

---

[4] The existence of conceptual objects implies the existence of intentional facts, which are non-material facts as they depend on mental states, more precisely, our mental capacity for representing (García-Carpintero, 1996, pp. 57-58).

[5] This distinguishes social norms, grounded in habitual behavior, from Lewisian conventions based purely on coordination incentives. While conventions may arise from strategic interaction without normative attitudes, social norms require participants to hold normative expectations —beliefs about what ought to be done, not merely what others are likely to do. *See* Bicchieri, 2006, pp. 11-28.

[6] As González Lagier explains, the truth of many of our statements depends on the conceptual network we employ for understanding the world. That is, truth does not rely solely on how reality is, but also on shared schemes for understanding it (2022, pp. 37-38).

Scarce attention to custom as a theoretical problem and overreliance on written sources have led not only to a limited understanding of international law but also to conceptually unsound explanations of customary legal norms, which fail to distinguish between ontological questions (what makes a norm exist) and epistemological ones (how we know it exists), and hence necessarily look to account for customary legal norms as if they were the product of prescriptive discourse[7]. This would be unsatisfactory for norms of customary sources, such as social norms, and it is unsatisfactory to claim that legal norms are ontologically distinct from norms elsewhere. The difference between a legal norm and other kinds of norms in the social world is not ontological; instead, it depends on the conditions employed to assign them to one system or another (Mattei-Gentili, 2024, p. 131).

To prove the existence of a customary norm, we must look beyond the law's doctrine to understand what it requires. To grasp the conceptual implications of norms derived from customary sources, it is first convenient to elucidate how statutory norms are produced and identified. The latter are deliberately enacted by a recognized authority and recorded in writing, whether as legislation, case collections, or precedential rulings. Contemporary legal theory is inclined to agree that, in the face of a controversy, a judge must exercise interpretation to identify the norm from the dispositions in the written legal sources. Thus, it is said, interpretation is to identify or assign a meaning to a text[8]. Conversely, customary norms are deontic meanings interpreted from recurrent and concurrent sets of social facts of a kind among members of a group. Hence, their identification involves two interpretative steps: one to

[7] For instance, Guastini has stated that customary norms do not exist unless they are formulated in a language (2018). Similar ideas are held by international law scholars, such as Markus Beham, in his «Schrödinger's custom theory.» He considers that custom (i.e., a customary norm) «[…] forms only in present, once it is invoked and an observer is introduced,» where for *observer* we must understand a court (Beham, 2022, p. 91). This is a misconception of Schrödinger's experiment and its lesson about the relationship between facts, objects, and knowledge. Schrödinger aimed to challenge the Copenhagen Interpretation, which holds that quantum systems have no definite reality until they are measured, contradicting the idea that facts have inherent properties. Schrödinger emphasized the nonsensicality of this view, noting that logically, the cat is either alive or dead, and uncertainty doesn't negate logic. We must distinguish between logical conclusions and our lack of knowledge until the box is opened. The same applies to customary norms.

[8] Within legal theory, there is a tradition of debate on whether interpretation is a creative or a cognitive activity. Among realist-influenced scholars, interpretation is seen as unavoidably creative. I do not fully endorse a cognitive approach, but radical legal realism is untenable. If interpretation always involves creativity, then legislation and language, as rule-based practices, become nonsense (*see* Navarro, 2005, pp. 121-122). Following Hart's lessons, there are cases in which the identification of norms is so straightforward, given conventional language usage and the conceptual framework we share, that interpretation involves a cognitive activity, or at least a highly cognitive one. These would be clear cases. However, in other cases, identifying a norm isn't obvious, and discretion is needed to resolve the question (Hart, 1994, pp. 128-129). Hart reminds us that discretion isn't the same as arbitrariness. When interpreting texts or facts, the interpreter's choice among multiple meanings is constrained by factors such as language, conceptual tools, and system rules (Hart, 2013).

determine that a chain of actions over time is customary[9], and another to infer that those customary actions involve a (social) norm.

Consequently, as deontological meanings, norms are markers about what is mandatory, prohibited, or permitted in a particular context[10]. The notion that norms are meanings is more straightforward to grasp when an authority employs language to communicate its intentions regarding what its addressees must do. But this scheme is not adequate for customary norms. Certainly, we can produce norms through agreements, which is assumed by the *pacta sunt servanda* principle. However, contracts are also the product of deliberate negotiation between parties, reflecting their intent to comply with the stipulated rules. This contrasts with the notion of custom, particularly social custom, which implies a recurrent behavior that is spontaneous (not deliberate) and that does not depend on linguistic means to exist[11]. Custom, and whatever custom brings about, exists because of recurrent practice, not because of language. Hence, we need an explanation of how recurrent behaviors can produce meanings, and how those meanings can become normative, *i.e.*, norms.

3.1. The Conventionalist Framework of Expectations

Spontaneous conventions[12] are how we produce social meanings in the absence of negotiated deliberation; through recurrent behaviors, they function as tacit agreements. In this sense, it said, language is a paradigm of a conventional product. David Lewis showed that we can achieve conventional outcomes without explicit agreements when repeatedly facing situations that require coordination to solve problems. We reach these solutions by examining each other's past behavior and inferring that, in similar situations, everyone will probably act as they have before. Others who wish to coordinate with us will also look at our past behavior in similar situations, inferring that we will likely behave as we have in the past.

---

[9] Vernengo stresses that customary facts are not material facts and, strictly speaking, are not mere empirical facts, but metaphysical abstract predicates. Therefore, identifying a particular action with the custom it displays involves an interpretative action (1976, p. 446). Also, Canale, 2008, p. 114.

[10] The idea that norms are meanings is a staple thesis in analytical legal theory for written sources, which tends to stress that legal norms are the meanings associated with, or interpreted from, a legal provision. In the case of customary norms, norms are meanings associated with certain social practices (Guastini, 2011, pp. 8-11). The idea that norms are meanings, language-independent (not language-indifferent), has been termed the *hyletic conception of norms* and explained by Alchourrón & Bulygin (1981, pp. 96-97). For an analysis of how the hyletic conception's ontological coordinates apply to customary legal norms, *see* Mattei-Gentili, 2020a. For a thorough exposition of the thesis that norms are meanings and what it involves for legal theory, *see* Mattei-Gentili, 2024.

[11] However, communicative resources, in a broad sense, are a requisite to produce them.

[12] I want to stress that spontaneity exemplifies conventions, especially as products of habitual social behavior, but we can also create conventional products through planning, negotiating, and expressive compromise (Celano, 2010: 212).

Lewis was aware that reaching these solutions is not easy; after all, we are inferring and adapting our behavior as we expect others to behave, so there will be a trial-and-error process. But once we reach a solution, once our behaviors converge to coordinate us onto a common path of action that solves the problem we face, we will praise that solution; it will be a common, *salient* option, and we will probably repeat it in future situations when we face similar problems. When the kind of situation and the solution become habitual, we will have formed expectations about each other's behavior: the expectation that the other will act in a certain way, so I must act in a specific manner in response, and vice versa. This means that our preference for behaving in accordance with the solution is partly conditional on others' adherence to it. The solution is common knowledge now, and we praise it for allowing us to coordinate. In this sense, the convention that is the solution acquires a normative impress; we represent it as what we ought to do in a given situation. In this sense, Lewis considered conventions to be species of norms: «Any convention is, by definition, a norm which there is some presumption that one ought to conform to.» (Lewis, 2002, p. 99).

Lewis' scheme of conventions as norms —as tacit rules[13] — is plausible for explaining a range of habitual social behaviors, but, for this reason, it seems at odds with the notion of *norm* we commonly conceive and that we are trying to figure out. In this explanation, conventions like fads are indistinguishable from more robust conventional social norms, such as promises. Lewis vaguely recalls that normativity is socially enforced by signs such as unfavorable responses from others, since failing to conform to a social norm involves going against the settled expectations of others, who can plausibly infer that one knowingly acted contrary to the convention[14].

Conventions are contingent; we could have chosen a different solution that also solved our coordination problem. That does not mean we are indifferent to them. Indeed, part of the reason we follow conventions is that others do too, which helps us coordinate in strategic social actions. However, this is not necessarily the only reason we adhere to these conventions. Agents tend to follow conventions as social rules and enforce them through adverse responses because they believe that some of these social rules reflect values beyond mere coordination (Chwe, 2001, chap 1).

Cristina Bicchieri has enhanced the framework towards an explanation of social norms[15], making the punctual distinction between conventional products like mere

[13] They do not need to be uttered to exist (Lewis, 2002, pp. 105-106).

[14] Brandom also assumes that normativity is the byproduct of the interplay among participants in a practice in which agents constantly make their moves and assess each other's actions, thereby emitting responses as signs of those assessments. He calls those responses «sanctions,» but we need not necessarily equate them with physical punishment or harm. Sanctions are merely the actions agents take to alter the normative status and attitudes toward actions (Brandom, 1994, p. 42).

[15] Bicchieri is not the only philosopher proposing an account of social norms. Diverse thinkers have engaged with the subject, but mainly as a byproduct in the scope of explaining wider phenomena like, for instance, the possibility of meaning (Gibbard, Blackburn, Brandom), the production of social reality (Searle), how rules operate and give sense to a social reality (Wittgenstein), the production

customs, what she calls «descriptive norms», and «social norms», and observes that the latter do not arise precisely from a coordination game, but from a mixed-motive game where once a solution is reached, turns into a coordination game (Bicchieri, 2006, p. 38)[16]. Also, social norms are equivalent to our understanding of customary norms (Postema, 2007, p. 286)[17]. They carry the deontic properties of obligation, prohibition, and permission, involve recurrent, convergent behaviors over time, and have unclear emergence times. If conventions imply an empirical expectation that, in certain situations, specific types of actions or outcomes[18] are held and will most probably be held, social norms arise when the element of normative expectations comes into play.

3.2. Normative Expectations

As Celano stresses, for a customary norm to exist, behavior must not only be repeated and considered binding, but also consistently observed, as it is regarded as such by the population (2010: 178). The idea is that observance of behavior is reinforced by the sense of duty in force because one believes that one ought to do so, primarily because others believe that one must do so, and so on. That is, there is a sense of expectation about each other's behavior that is not merely empirical.

Normative expectations evolve from empirical ones, introducing the *ought-factor*. That is, they not only predict that specific outcomes produced in the past will happen again, but also that they *ought* to occur. That means, when a certain social behavior stands for a social norm, we are not only aware of the informational insight about

---

of institutions (Weber, Searle, Tuomela, Hart, Weinberger, Marmor, Guala), to explain conventions (Lewis, Gilbert), to explain how diverse schemes of strategic interaction can produce rules (Ullmann-Margalit), to explore the possibilities of a deontic logic (von Wright), etc. There are meaningful insights from these approaches, and whenever useful, I will refer to them as Bicchieri has also done. However, the main aim of Bicchieri's account is social norms; thus, it provides a robust but straightforward insight that can be useful for explaining the sort of questions we face in understanding customary norms. An alternative to Bicchieri's account is the theory advanced by Brennan, Eriksson, Goodin, and Southwood. However, I find it unconvincing for several reasons: it dismisses a conventionalist framework without properly engaging with Lewis' account — reducing preferences in social contexts to mere personal desires; it misconstrues the ontological status of norms, treating them as vehicles of social meaning rather than meanings in themselves; and it ultimately offers a circular definition of norms as «accepted rules or normative principles.» (Brennan, Eriksson, Goodin & Southwood, 2013, p. 2).

[16] In game theory, a mixed-motive game involves parties with both cooperative and competitive goals. When repeated, and parties see cooperation as more beneficial than selfishness, the game becomes cooperative, favoring coordination over other actions. This creates an environment conducive to the formation of conventions and, more specifically, social norms.

[17] This is in line with the common understanding of international law scholarship of Customary International Law as grounded in a particular social reality (d'Aspremont, 2022, pp. 29-30).

[18] Not only can public actions produce social outcomes like conventions and social norms. Some private actions produce publicly notorious outcomes, hence making propitious to produce social expectations.

past behaviors, but we also believe that we *ought* to act by those past behaviors, and, somehow, we are aware and believe that others believe we *ought* to abide by those past behaviors.

Empirical expectations and normative expectations involve different cognitive attitudes. The former are first-order beliefs about what others will do: «I expect most people will perform action X in situation S.» The latter involves second-order beliefs about what others think people *ought* to do. More precisely, a normative expectation has two components: (a) a *normative belief*—«I believe others think people should perform X in S»—and (b) a *normative expectation proper*—«I believe others expect *me* to believe people should perform X in S.» This recursive structure creates conditional normative preferences: individuals prefer to conform to a social norm *on condition that* they believe others share both the normative belief and the expectation that others hold that belief.

This second-order structure explains a range of puzzling features of social norms. First, it explains why social norms can exist even when participants privately disagree with them — what matters is not individual endorsement but shared expectations about what the group normatively expects. Second, it explains norm enforcement: criticism of norm violations appeals not to the critic's personal preferences but to what «we» (the group) expect of each other. Third, and crucially for customary law, it is a response to the apparent circularity in *opinio juris* understood as a mental constitutive element of a customary norm. The belief that «there is an obligation to perform X» need not precede practice for practice to manifest a normative mental state that could be called «*opinio juris*.» Rather, what is required is the normative expectation structure: States engage in practice X while believing that other States expect them to follow X and expect them to hold that expectation. This is not a belief that X is *already* law, but a network of conditional normative expectations that *constitutes* the social norm.

For evidentiary purposes, this conceptual clarification is critical. Courts seeking to prove *opinio juris* need not identify States' beliefs that conduct is legally obligatory (which would be circular) [19]. Rather, they must identify evidence of normative expectations: manifestations that States believe others normatively expect certain conduct and that others are expected to share that expectation. Such evidence includes not only explicit statements invoking obligation, but also patterns of criticism (indicating normative expectations), justifications for conduct (appealing to what «the

[19] For instance, this is a conceptual mistake in Conclusion 9 of the United Nations' International Law Commission's *Draft conclusions on the identification of customary international law*, which emphasized that practice must be undertaken with a sense of legal obligation when determining the *opinio juris* existence. Part of the problem with inquiry into the existence of customary legal norms is that, when discussing the *opinio juris* element, jurists focus on finding the basis of legal membership rather than on identifying a necessary condition for the existence of such norms. One inquiry is proper to legal theory, while the other is conceptual-ontological. (Celano, 2010: 179-180).

community of States» expects)[20], and responses to perceived violations (expressing disapproval rather than merely strategic responses to unwelcome behavior). The distinction between empirical and normative expectations thus provides a framework for determining what evidence is relevant to proving customary norms. Also, from the standpoint of a factfinder, the reasoning process is dual, giving rise to inductive reasoning to acknowledge empirical expectations and to abductive reasoning to acknowledge normative expectations. Hence, recurrent and convergent States' practices permit inductive inference of a common custom (*usus*). But States criticize deviations, manifest discord to uncanny behaviors, justify their conduct by appealing to what is expected by the community — they don't simply act, but also react to others' actions. These reactions are striking and call for explanation. Therefore, among the diverse available hypotheses (such as protecting self-interests or political prudence), the hypothesis that States manifest these reactions because of a framework of normative expectations (*opinio juris*) provides a more comprehensive and coherent explanation of these phenomena. Accordingly, the existence of a customary norm is *abductively inferred* as the best explanation of the observed manifestations of normative expectations.

Now, there is no indisputable feature that explains why a particular pattern of behavior acquires normative expectations and becomes a social norm (Bicchieri, 2006, pp. 41-44). A pattern that in society *A* is a descriptive norm in another society *B* can be a social norm. And what once was merely a descriptive norm in society *A* can become a social norm in T1, then return to a descriptive norm in T2[21]. The reasons some convergent practices produce social norms are infinite, and to understand why a social pattern of behavior becomes a social norm, one should engage in an ethnographic inquiry into the factors that forge a society's idiosyncrasy. From the conceptual enterprise of what could be deemed as an exercise of armchair descriptive sociology[22], one can only stress features of conditional social expectations and social preferences to identify normative expectations that give rise to social norms[23].

[20] As Bulygin has emphasized, norms cannot only be employed for sanctioning or criticizing deviation, but also for justifying our actions (2005, p. 32). Similarly, Postema's discursive account emphasizes the relevance of reason-giving among participants in a reciprocal normative practice, among which justification is fundamental (2012, pp. 732-734).

[21] This phenomenon has been acknowledged in international law by Kunz, noticing that what once was a norm of *courtoisie internationale* can become a norm of customary general international law by the later addition of an *opinio juris* to the practice of States (1953, p. 667).

[22] The idea that explaining social institutions and products involves a kind of speculative sociological exercise is held among diverse philosophers (Hart, 1994, p. v; Ullmann-Margalit, 1977, p. vii).

[23] As Bicchieri has emphasized: «The triad of empirical and normative expectations and conditional preferences is what, in my view, *defines* social norms.» (2014, p. 226).

3.3. Normative Attitudes Without Moral Endorsement

Conventional products, like social norms, imply some form of assessment of most of the population. However, it should not be interpreted as a strict moral judgment, which leads to a moral norm. One does not follow moral norms simply because others do. Conventional products obey adaptive behavior. In conditions of uncertainty, we tend to imitate others, others we deem successful or, at least, deem that they know what they are doing, what is «correct» in such a situation (Bicchieri, 2017, p. 23). This means that when a convention is in place, our actions are interdependent, that is, they involve a social conditional preference, *i.e.,* we do it partly because others do it too. If others started changing their behavior, one would do the same, not only because it would be evidence that social assessments have changed, but also because one desires to coordinate one's actions with one's peers. Therefore, if we value conventions — and as such, social norms — in a relevant sense, it is because they allow us to coordinate with others in social enterprises. Indeed, for social norms to emerge, what the group assumes lies behind the convention's coordinative solution must be positively valued. But this is a delicate matter; some coordinative functions that conventions and social norms preserve are notorious, and it's easy to understand why people value the norm and therefore punish defectors. Traffic rules are clear examples, but one can think of others, such as tradesmen's conventions or even the paradigmatic relevance of the institution of the promise. But the value behind other social norms is not that obvious.

Many people may believe that a certain norm is less effective at achieving social goals than another conceivable one. It is possible that most individuals despise the current norm but are afraid to express it in the face of potential sanctions — they believe others morally endorse the norm, so they remain silent[24]. Therefore, the social norm remains in place because, although people regard it as objectionable or ineffective, they also consider it valid and enforce it on others. This structure is compatible with what, in legal theory, Hart described as the internal point of view: adopting a normative stance toward a pattern of behavior (Hart, 1994, p. 90. As Shapiro suggests, it should not be regarded as an insider's point of view — the stance of someone in the community whose only interest in the rules is to avoid sanctions — but as a normative attitude towards the pattern of behavior that instantiates the rules. And somehow, this normative attitude involves a critical-reflexive attitude on the part of the agents, enabling them to recognize and accept such patterns as valid rules of society (Shapiro, 2006, pp. 1159-1165).

[24] This is what Bicchieri calls *pluralistic ignorance* (2017, p. 42). Also, Sunstein, 2019, pp. 5-6.

4. CUSTOMARY *LAW*. A THEORETICAL SUBJECT

Legal theory has an epistemological scope, aiming to stipulate the truth conditions for statements about legal duties and rights and to provide criteria for identifying the legal norms that support those duties and rights (Moreso, Navarro & Redondo, 2002, p. 7). Therefore, I adopt a positivist outlook, aligned with the social philosophy that conceives of norms as products of social practices (social facts), emphasizing the importance of a value-free description of law[25].

As emphasized, there is no ontological feature of legal norms that distinguishes them from norms in other realms of social life. Therefore, there is no intrinsic feature in customary legal norms that distinguishes them from other customary norms (*i.e.,* social norms). What, then, makes a norm pertain to the domain of law? First, Kelsen suggested that there are no isolated legal norms. Whenever there is a legal norm, there is at least another norm related to it, underscoring the relevance of the notion of system for understanding law. In a further step, Hart established that what makes a norm *legal* is its membership in a system, *i.e.*, a legal system. Thus, law is a system composed of two types of norms: primary and secondary. Primary norms are basic rules of conduct, while secondary norms concern the operation and structure of the legal system, providing mechanisms for producing, adopting, changing, and enforcing legal norms. Consequently, a norm is part of the legal system if it has been created or recognized by a secondary norm; it can be a rule of change, a rule of adjudication, or a rule of recognition.

Hence, *prima facie*, a customary norm can become part of a legal system if it has been adopted or recognized by a previously existing legal norm. The rules of recognition can recognize several customary norms as part of the law; the rules of change can adopt customary norms, and the mechanisms of legal adjudication can also incorporate them (Mattei-Gentili, 2026, pp. 67-68).

In legal systems, which mostly rely on statutory law, a legal provision can recognize certain customary norms as part of national law. For instance, Article 2, A. I., of the Mexican Constitution guarantees and recognizes the customary normative systems of indigenous peoples as valid law under which they have the right to organize their communal life. Therefore, it is plausible to assume that statutory law already confers *legal* character on those customary norms. The issue is more complicated in systems that do not depend on statutory law or where statutory law does not clearly specify which customs are sources of law[26]. This is the case with international law.

[25] It is also worth noting that international law scholars like Sender and Wood recognize that the approach of customary law is grounded on facts and not on preconceptions about what law ought to be (2004, p. 79). Hence, the approach can be seen as a test of positive law.

[26] For instance, the Nigerian Constitution implicitly acknowledges custom as a source of law, even establishing Customary Courts of Appeal for each State that needs them. However, there is no legal

To determine whether a customary norm belongs to the system of international law, custom itself is the main available source.

Within Hart's theory of legal systems, the rule of recognition is itself a customary norm; more specifically, it is the outcome of the practice of legal officials in identifying ultimate valid law. If this were accurate, the foundation of any legal system would be customary law, meaning international law would be based on the same principles as national legal systems, including those that primarily depend on legislation. That would mean customary legal norms underpin further legal norms of change and adjudication. However, this offers little insight into the dynamics of customary norms within legal systems. For such matters, Bentham suggested distinguishing between customary norms *in foro* and customary norms *in pays* (Bentham, 2010, pp. 234-235). The former are the outcome of legal officials' practice, while the latter are the outcome of citizens' practice. Since a customary norm *in foro* (the rule of recognition) is the ultimate basis for the validity of any legal norm, the validity of further customary norms as legal norms depends on the former[27]. Hence, the absence of a formal instance of recognition of customary norms *in pays* as members of the law, the only instance left for such recognition is judicial practice and its customary norms *in foro* (Mattei-Gentili, 2026, pp. 91-93) Let us explore what this demands for the epistemic purposes of this essay.

First, it is important to stress that, in the face of controversy, proof of customary law *in foro* is excluded since its validity is assumed to be granted. For international law, this means that only customary norms *in pays* — norms form the customary practice of States — are subject to the evidentiary legal procedure. However, the fact that judicial practice and judicial decisions ultimately determine what is recognized as customary law *in pays* leaves room for epistemic error. It leaves open the possibility of producing «customary» legal norms that are not the product of States' practice. Consequently, it is conceivable to have genuine customary legal norms *in pays* when judicial practice succeeds epistemically in identifying actual customary norms among the practice of the State members of the international community, and spurious customary legal norms *in pays* when judicial practice epistemically errs by determining as valid an alleged customary norm that is not supported by the actual practice of members of the international community. Both would be international legal norms that derive their legal validity from the international rule of recognition, but only the former genuinely reflects the actual customary practice of States. The latter are judicially created norms.

Among international law scholars, it is common to conceive of customary international law as disguised judge-made law (*see* Talmon, 2015). The assertion is too

provision in the constitutional text specifying which customary norms should be considered part of Nigerian law, as opposed to those that are not.

[27] Recognition of the custom of the courts as a source of law is as old as, at least, the Roman Empire. As Bederman notices, in § 1.3.58 of the *Digest*, Callistratus exposes that «in cases of ambiguity arising from statute law, statutory force ought to be ascribed to custom or to the authority of an unbroken line of similar judicial decisions.» (Bederman, 2010, p. 19).

categorical, but courts can err in identifying customary norms. This possibility allows for creative decisions that may create general legal norms if judges uphold them as legally valid[28]. However, the theoretical lesson relevant for evidentiary purposes lies in what makes a customary norm *in pays* a customary *legal* norm.

Much of the evidentiary challenge lies in identifying the subjective and psychological elements underlying the recurring practice. And while it is true that a psychological element is unavoidable for producing customary norms, the peculiarity of that psychological component is not what makes a customary norm *in pays* a legal norm — the same is not necessarily true for customary norms *in foro*. One could argue that, if the rule of recognition of international law accepts customary norms from States' practice as part of the system, they are, by default, considered customary legal norms *in pays*. What makes customary norms *in pays* a part of the law is their compliance with the criteria of membership to a legal system. One of these criteria is recognition by international courts. That recognition, to be epistemologically justified, requires that courts examine the normative social beliefs that accompany the concurrent States' practice, but they are not conceptually required to also find that States believe that such practice involves *a legal* obligation. The court's duty consists of acknowledging the conceptual elements comprising customary norms and applying the rules of epistemic reasoning to prove their existence.

5. CONCEPTUAL CLARIFICATION FOR INTERNATIONAL LAW DOCTRINE

Some of the problems international courts face in determining the existence of customary norms are not strictly epistemically, but stem from conceptual confusion regarding legal doctrine. One involves the two-element formula, particularly understanding the *opinio juris* element. The other concerns the practice's self-understanding of its outcomes in determining the existence of the components of customary norms.

There are several discussions and little agreement on how to characterize the elements of the two-element formula traditionally recognized in legal doctrine. Nonetheless, the formula remains regarded as the best blueprint for proving the existence

[28] Hans Kelsen regarded judicial rulings as individualized legal norms. Eugenio Bulygin notes that if that were the case, each time a judge makes a ruling, the legal system, as a normative system, changes, posing difficulties for the idea of law's continuity. Hence, he suggests legal norms are only those criteria that have been generally upheld (Bulygin, 1991, p. 262). This means that a new interpretation of a legal disposition does not add a new norm to the system; for that to happen, the interpretation must become generally employed by other judges — there must be a customary practice among judges of recognizing it as valid law. I consider this a hint at the rule of recognition and customary norms *in foro*. However, in international law, there are few courts, and cases are significantly scarcer. Each international resolution tends to attract significant attention and is often presumed to entail the immediate production of a general international legal norm. *See* Tams, 2015.

of customary legal norms, and international law is no exception[29]. Since international courts will likely continue to employ the formula, it is appropriate to resort to rational reconstruction to clarify its components and the process of identification.

5.1. The Relevance of the *Opinio Juris*

There are two main lines of debate and proposals surrounding the two-element conception of customary legal norms. One involves removing the *opinio juris* as a constitutive element and, therefore, focusing exclusively on the less controversial component of the *usus*. The other comprises recognizing that some customary legal norms are composed exclusively of *usus*, while others are composed solely of *opinio juris*. However, these lines of argumentation cannot withstand conceptual analysis and scrutiny.

As Chiassoni emphasizes, although there are diverse arguments against *opinio juris*, the logical and epistemological arguments are decisive (2008, pp. 73-74). The former criticizes a vicious regress: *opinio juris* is a prerequisite for the creation of a customary norm, but at the time, the existence of *opinio juris* as a conviction to submit to a legal obligation implies an already constituted legal norm (Bobbio, 2010, p. 57)[30]. The solutions, as Bobbio suggests, are: considering the *opinio juris* as based on an error and, hence, customary law is grounded on a mistake[31], or admitting there is already a constituted norm before the manifestation of the conviction, but, consequently, reducing to irrelevance not only the *opinio juris* as constitutive component (Bobbio, 2010, p. 58), but also the *usus*, since the norm would already be formed, and practice merely involves a source of cognition of law, not of production (Chiassoni, 2008, p. 74). Conclusively, Bobbio considers that the only satisfactory solution is to get rid of the *opinio juris*. The epistemological argument holds that *opinio juris* must be eliminated from the characterization of customary legal norms, since mental states are not susceptible of being known, if at all, only conjecturally through external behavioral clues (Chiassoni, 2008, p. 75).

Elsewhere, I have explained why I am dissatisfied with accounts that dismiss mental features as key parts of customary norms (Mattei-Gentili, 2020b, pp. 91-

[29] Article 38. 1. b, of the Statute of the International Court of Justice even stresses that applying custom as a source of law involves «[…] evidence of a general practice accepted as law.» This could suggest that custom is not a source of law but merely proof of pre-existing (natural) law, thus distinguishing validity as membership from validity as applicability [*see* Bulygin 1991: 266 - 267]. The phrasing is best understood as a historical relic (the original wording is from 1920). Under a positivist framework, however, customary legal norms are valid because they meet the system's membership criteria. The text could thus be understood otherwise as considering customary international norms not as members of the international legal system, but merely as applicable legal norms, and general practice as the only evidence of their existence. However, the ICJ applies custom as law, not as mere evidence.

[30] Kelsen presents another version of the error theory (Kelsen, 1949, p. 114).

[31] Anyone convinced to comply with a legal obligation is mistaken. As Bederman notes, this line of critique of the *opinio juris* is as old as the Romans (2010, p. 20).

93). Also, I see no problem with customary norms being based on a logical mistake of judgment, which does not make them less customary nor less normative[32]. More importantly, as we have seen, social philosophers do not avoid discussing mental states in their explanations, but also highlight the importance of expectations and inferences about others' mental states in understanding social norms[33]. If this is so, it is because removing the mental features that accompany our practices would exclude conventions and social norms altogether. Conventions and norms are not material facts; they do not exist without a mental act by their creators. Now, it is true that we can never be certain about others' mental states; inferring a subject's mental state from behavioral cues is complex, error-prone, and susceptible to misinterpretation[34]. This means that inference to others' mental states is fallible and operates under irreducible epistemic uncertainty, but that does not mean we cannot achieve justified knowledge of them. Fallibility is a feature of all empirical knowledge, not a mark of mere conjecture. What matters is whether inferences are grounded in reliable indicators and subject to intersubjective scrutiny[35]. Not only are we prone to infer others' intentions in our everyday interactions with high success rates, but legal practice is no stranger to the evidentiary task of determining the existence of mental states, with *mens rea* serving as a paradigm in criminal law.

What we need is an explanation of what those mental states — which can be called *opinio juris* — encompass, given what we know about social norms. In this sense, the first issue to tackle is semantic. Bobbio suggests that, while there is no doubt that *opinio juris* indicates a mental state, it is not clear which sort. Doctrinal developments have suggested that it is sometimes understood as a conscience or conviction, less commonly as a belief or a feeling, and exceptionally as an intention or will. When adding «*sive necessitatis*» to the expression, it is unclear if the actions must be carried out of conviction of acting out of necessity, or properly out of legal necessity, or even as a conviction to avoid sanction (Bobbio, 2010, pp. 56-57). Finally, the term suggests that the practitioners' mental state makes any norm *law*, that is, part of the legal system.

---

[32] If we assume Hume's law as adequate, all our normative conceptual apparatuses are based on the logical mistake of deriving normative judgments from factual judgments. That does not make things like language or promises, for instance, less real or less useful. As Bicchieri has pointed out, «[…] personal as well as historical evidence tells us that we are readily victims of this "naturalistic fallacy".» (Bicchieri, 2006, p. 43).

[33] In a similar vein, dismissing the error theory critique, Canale stresses that any belief capable of motivating the qualified repetition of a behavior satisfies the requirement of the *opinio juris*. Therefore, error theory turns into a mere *a posteriori* psychological conjecture (Canale, 2008, p. 120).

[34] This is about the *problem of other minds*. I see others' actions, which suggest they have certain mental states, but I can never be sure of their actual mental states or motivations.

[35] González Lagier's *principle of minimum rationality* (PRM) is a suitable epistemic tool for grounding inferences and enabling intersubjective scrutiny. It states that when an agent acts intentionally, she performs the action she believes best achieves her end. This principle helps assess intentions by reconstructing the agent's practical reasoning within the relevant context. González Lagier, 2005, pp. 207-216.

The task of identifying customary norms is bound to be speculative to some extent, and demanding that everyone involved share the same precise mental content means we will find nothing. It is conceptually and epistemologically too burdensome. We don't need to all have the exact same idea in mind to share conceptual items such as norms. About a social norm, some may have the conviction that the norm is needed, some others may despise the norm but believe it is valid and obey it nevertheless, some may consider the norm as necessary for a social aim but some other may not have ever thought much its utility; moreover, that same population may even disagree about the proper linguistic formulation of the norm (Bentham, 2010, p. 161; Gardner, 2012, p. 67). Much of this involves controversy, but that does not mean that a customary norm is not in force (Mattei-Gentili, 2026, pp. 244-245). Norms are not language-dependent, and social norms can serve diverse functions, but that is not necessarily the reason they exist. So, discussions about their formulation might merely involve linguistic disputes, and debates about their social purposes merely involve projections about diverse political and moral appreciations. Despite the lack of agreement on either question, the community continues to exhibit the same pattern of behavior and manifests the same normative attitudes towards that practice; that is, they share the social norm.

Patterns of behavior identifiable through a concurring social practice held over time, and normative attitudes attached to that conditioned practice, not only constitute social norms but also serve as clues for social scientists as evidence of existing social norms. Conditional preferences and normative attitudes are indicators of normative expectations shared within the population, and these expectations can plausibly be characterized doctrinally as *opinio juris* (Celano, 2010: 222). Nevertheless, as clarified, the normative expectations do not necessarily entail believing they are legal duties.

5.2. The Issue of Inadequate Justifications

Courts are not ethnographers nor social philosophers. While part of the international community, they are not involved in creating customary norms *in pays*. That explains some of the issues they face, but it does not mean they are impeded from acknowledging the existence of customary norms *in pays*. Some of the gaffes are not strictly epistemological, but argumentative. That is, it is not uncanny that international courts are untidy in justifying their decisions when determining the existence of customary norms. In other words, courts do not make a neat exposition of their reasoning in assessing evidence and determining the existence of customary norms. For instance, it has been alleged that the International Court of Justice (hereinafter, the ICJ) has determined the existence of a customary norm but failed to acknowledge consistent State practice in cases such as *Nicaragua v. United States*. It has also established the existence of a customary norm without resorting to *opinio juris* in *The Nottebohm Case* (Kirgis, 1987, pp. 147-149). This has led to confusion among some

scholars. In the conceptual field, Roberts has concluded that there are two kinds of customary international norms: traditional ones, which are determined by induction from the accreditation of consistent State practice, and modern ones, which are determined deductively from verification of an *opinio juris* among States (Roberts, 2001, pp. 757-758). In the methodological field, Kirgis has proposed the sliding-scale model, under which, depending on circumstances, the courts can determine the existence either from evidence of *opinio juris* or from evidence of consistent State practice.

Customary norms, as social norms, are ontologically consistent: all of them are composed of a convergent and repeated practice and a shared mental state from their beholders[36]. Instead, Robert's account suggests that it is courts that create the norms of customary international law, justifying their actions based on either *usus* or *opinio juris*. For its part, Kirgis does not notice that courts' methods or justifying arguments may be deficient. Instead, he justifies the courts' practice as adequate because of the values or the goods at stake. That is, methodologically, a court may not do all that is needed to prove the existence of a customary norm and, erroneously, determine its existence based on evidence of only one component. Or it may be that a court did find evidence of both components (and, hence, found a customary norm), but in the argumentation that justifies its decision, the court fails to refer to one of the elements.

Genuine customary norms *in pays* are not judge-created. That international courts in the past tended to accentuate evidence on State practice to prove the existence of customary norms does not mean that traditional customary norms are the result of consistent practice followed by the States ignoring *opinio juris*, while modern customary norms are identified mainly by State statements that emphasize the existence of an *opinio juris,* neglecting States' practice. The change in the approach to evidence regarding the elements does not change the nature of customary norms (Mattei-Gentili, 2026, pp. 106-107). If a court fails to find one element but still determines the existence of a customary norm, it is an erroneous ruling. Should a court ascertain the presence of both constitutive elements but fail to articulate its reasoning for each explicitly, the decision's justification is deficient.

[36] In international law, Kelsen notes, States' behavior reflects that of the individuals representing them (Kelsen, 1949, pp. 191-192, 364). This insight is essentially adequate, but it must be complemented: the alleged «mental state» of a State is best understood as a shorthand for the resulting pattern of actions of the individuals who have represented it over time. These actions are guided by internal institutional rules that shape the representative role —for instance, the Chancellor, the head of State, etc. Some States are consistently more proactive than others, and this is observable across successive representatives. Strictly speaking, however, nonphysical agents like States do not possess mental states. When referring to a State's mental state, we refer to its representatives' mental states, which are influenced by institutional rules. Individual agents retain their own motivations and responsibility for their actions, including violations of international law that are inconsistent with the pattern of action of the State they represent. On epistemic matters regarding State representatives and their actions. Collins & Lawford-Smith, 2016, pp. 151-152, 159-159.

6. THE EVIDENTIARY METHOD AND APPLICATION

Proving customary international norms faces fewer formal constraints than evidentiary procedures in national legal systems. Hence, there are fewer epistemic obstacles. In this regard, the United Nations International Law Commission has left the burden of proof open. While in national legal systems, parties bear the burden of proving the existence of customary norms[37], in international law, the *jura novit curia* principle places this responsibility on courts, though parties may also provide evidence (Sender & Wood, 2024, p. 98). Nonetheless, this procedural difference does not alter the fundamental epistemological task. In any case, courts must evaluate evidence to determine whether a customary norm exists, whether the evidence comes from the parties or the court's own inquiry.

The primary doctrinal constraint international courts face is that usus and *opinio juris* must be separately proved, though how this is to be done remains unclear. In what follows, I propose an appropriate interpretation of this. Next, I examine the epistemic challenges courts face and propose an interpretative approach to resolve them. Lastly, I briefly address a largely ignored issue: the lack of evidentiary standards.

6.1. Inferring Mental States from Repeated Actions

Sender and Wood note that the existence of *opinio juris* cannot simply be deduced from the existence of an usus; acting in a particular manner does not prove anything of a legal nature[38], and, because acceptance as law must be sought with regard not only to those taking part in the relevant practice, but also in those in a condition to react to it (Sender & Wood, 2024, pp. 89-90). This idea has been reinforced by the International Law Commission of the United Nations, which, in paragraph 2 of Conclusion 3 of the *Draft conclusions on the identification of customary international law* of 2018, states: «Each of the two constituent elements is to be separately ascertained. This requires an assessment of evidence for each element.»

As explained, when dealing with customs *in pays*, acting in a particular manner can be an input for inferring the existence of normative expectations, constituting evidence for the existence of a customary norm. Consequently, external actions are the only means available for inferring mental states such as normative expectations. The proof of an *opinio juris* cannot be separated from evidence of the *usus*. The condition of separately assessing each element requires the court to provide specific reasons for each constituent, not different evidentiary materials, ensuring a separate argumentative justification.

---

[37] For instance, article 3.1 of the Spanish Civil Code requires parties to provide evidence of the existence of a legal custom to enforce it.

[38] This has been stressed by the ICJ in the *North Sea Continental Shelf* case.

The emphasis in this necessity is comprehensible since the ICJ has not always been meticulous in manifesting either the existence of an *opinio* or an *usus* in the determination of customary norms, raising doubts either if it was merely a defective labor, if the court found the uncanny case of customary norms composed by only one element, or if it acted arbitrarily and determined there was a customary norm where there was none. Although there are no exceptional cases of customary norms composed of a single element, it is still possible that some decisions are either wrong or arbitrary; however, some cases of determining customary norms involve enthymematic arguments in which not all premises are explicitly stated, raising questions about the logical validity of the conclusions.

Occasionally, the ICJ assumes the existence of *opinio juris* among participants when proving a persistent, convergent practice as *usus*, without mentioning it in its resolution. More controversially, commentators have lamented that the ICJ has sometimes only referred to the constatation of an *opinio juris* by the parties without reference to a *usus* (*see* D'Amato, 1987). Can a court legitimately — on epistemologically sound grounds — verify that someone has a mental state without reference to any external action? Since telepathic powers are excluded, the answer is negative. We can infer others' mental states only from what they do, including what they say. Declarations are material external actions. This seems to escape commentators who exclude declarations as instances of *usus*. However, as speech acts, declarations can and do create and perpetuate genuine customary norms[39].

The inquiry for the *opinio juris* requires an interpretive approach that does not ignore that State practice (*usus*) is situated in a contextual matrix of intersubjective expectations, through which its normative significance is constituted and maintained[40]. It must also be flexible and allow diverse concrete actions to be included within the abstract type of action demanded by the norm[41]. That includes actions

[39] In this sense, Brandom proposes that normative language is employed to make explicit support for a *pattern* of material practical interferences. Those practical inferences can range from mere conventions to customary norms that, up to that point, had been implicit in a group's convergent practice (Brandom, 2000, pp. 89-90). This is also somehow presupposed by Searle's regulative rules, which govern preexisting activity and which we come to know through *speech acts* (Searle, 1969, pp. 33-34,). Postema also emphasizes that action and word work together in the formation of customary norms (Postema, 2007, p. 29). In this regard, it is worth noting that some International Law scholars have argued in favor of considering declarations as components of State practice (Akehurst, 1975, p. 53).

[40] Perhaps this is what d'Aspremont intends when advocating that while identification of customary norms is an *in abstracto* question, evidence and ascertainment of the norm is an *in concreto* process (d'Aspremont, 2011, p. 151).

[41] For instance, if the norm requires communicating intentions to another party, this can be done through diverse concrete actions, such as formal diplomatic letters, press conferences, or even legislative actions. In this line, the need for flexibility in interpreting evidence to ascertain facts is noted in the second commentary to the third conclusion of the *Draft,* and in its fifth conclusion, which states that legislative actions and judicial rulings are State actions to be considered when assessing conformity with a customary norm. As Bicchieri underlines, many norms admit variations in behavior, and the range of variation can be substantial (2017, p. 36). Hence, Conclusion 7 of the Draft states that variations in a

such as what parties declare, manifesting acceptance of a norm (as a critical reflexive attitude) that converges with the common actions and expectations of other parties in the contextual practice; that is, the practice of declarations displays the *opinio*. This means that declarations supplement the other actions that constitute the practice of States in supporting a customary norm like the one of non-intervention[42].

In a case like *Nicaragua v. United States*, declarations were actions in which the United States, along with other nations, converged on consolidating empirical expectations (*usus*) with normative expectations (*opinio juris*) in a matrix of conditional preferences. Thus, although the United States breached the non-intervention norm on diverse occasions, that did not mean it did not recognize the norm, nor that it did not participate in its creation and maintenance. The United States normally does not intervene in other nations' affairs[43]; moreover, it has recognized this norm in various declarations and has repudiated such intervention when other nations intervene in others' affairs. From the perspective of social norms, an individual can transgress a norm on diverse occasions, a norm they nonetheless acknowledge as valid, ordinarily comply with, and even actively enforce upon others. Focusing exclusively on actions of non-abidance would imply that no party would ever participate in the creation and maintenance of the norm, nor accept it. Therefore, courts must distinguish between acts of transgression, as in the United States in the *Nicaragua v. United States* case, and non-participation or even repudiation of the norm-practice, as in Germany in the *North Sea Continental Shelf* case[44].

Attaining the context contemplates being aware that the practice that raises the norm involves not one but diverse parties, and that there are other norms, a system of norms, that compose international customary law, that also give sense to the normative practice of the examined alleged customary norm. This is particularly relevant when assessing the existence of customary permissive norms. This is illustrated in the *Right of Passage over Indian Territory* case, where the ICJ determined that Portugal had a pacific right of passage over the Indian territory for private persons, functionaries, and merchandise to access the Portuguese enclaves. This meant Portuguese persons did not have to apply for a permit from the Indian authorities. India claimed Portugal had

---

State's practice must not be immediately considered to constitute a new custom; this must be assessed in light of the circumstances.

[42] Conclusion 10 of the Draft emphasizes this for the proof of *opinio juris*.

[43] The phrase is controversial since the United States has interfered in other nations' affairs multiple times. Therefore, contextualization is needed. What I mean is that, although the United States has unjustifiably interfered in the affairs of other nations on repeated occasions, it does not do so with most nations and does not do so consistently. Moreover, on most occasions, it has done so by trying to hide its actions, which, as will be argued, is a clue to the State's recognition of the existence and validity of the customary norm. That is, at least until 1986, the United States was no persistent objector to the non-intervention customary norm.

[44] In this case, Denmark and the Netherlands were pushing towards the creation of a customary practice to claim a customary norm. However, from the beginning, Germany not only abstained from participating in the practice but also rejected it.

no right and had been tolerant, a concession revocable at any time. The ICJ reasoned that a practice lasting over a hundred years suggested more than a concession; it was plausible to infer a permissive bilateral customary norm, as nations usually request permission to pass through foreign territories. That is, the permissive norm for Portugal was inferred not just from India's never objecting for over a century, but also from the fact that, in the common customary norms of the international community, there is a mandatory norm requiring permission to access other nations' territory[45].

6.2. Narrative Interpretative and Approach

Courts face challenges in determining whether customary norms exist when interpreting evidence. Not being exhaustive, I find the following. First, the epistemic distance involved in being a third-party alien to the practices that create and support the customary norm. Second, the fact that customary norms conceptually entail coordination and, *prima facie*, common self-interest for acting, hence raising the question: Why would anyone act against their own self-interest? Third, how to understand *sanctions* that aid in inferring the existence of a customary norm, if international law excludes unilateral sanctions?

The above questions concern the problematic issue of judicial discretion. But discretion is not arbitrariness. Epistemologists acknowledge that discretion is unavoidable when assessing evidence and *making decisions*, but this does not imply an epistemologically unjustified choice. Hence, the issue is: What can courts do to justify, on epistemic grounds, their rulings about the existence of customary norms?

A party's conduct has diverged from her usual behavior and community expectations. The key question is whether this deviation represents a mere departure from a custom or a violation of a norm. Brandom considers that norms are in the eye of the beholder, one embedded in the practice. Therefore, addressing a group's customary norms requires considering what it means to acknowledge them in practice (Brandom, 1994, p. 25). Once more, norms are distinguished from mere practices from the internal point of view[46]. This is a significant obstacle for courts as they are not firsthand actors in customary norms *in pays*[47]. Difficulties increase when considering

[45] Although Portugal argued that its right of passage also included armed troops and military equipment, the Court ruled that the right was limited to civilians, officials, and merchandise. Additionally, as Johnson points out, the Court once again failed to explicitly refer to evidence of an *opinio juris* (Johnson, 2021, p. 1184).

[46] Hart's internal point of view can be expanded using Brandom's explanation that norms can be conceived as a type of practical commitment shared within a group. Therefore, norms help explain why someone acts a certain way and make it understandable to others in the group. (Brandom, 2000, p. 93).

[47] I stress this because, in a sense, they are part of the international community, and their resolutions also shape the behavior of States, including behavior that can lead to the formation of customary norms. As Bentham notes, judges' actions can influence the population's behavior and produce social norms (Bentham, 1977).

that, conceptually, customary norms are the outcome of a coordination solution that serves the best interests of the parties. But actions of defection from the coordination solution are usually argued by the parties as supporting their actual self-interest, hence, trying to defeat the claim that such a pattern of conduct represents a genuine customary norm. The account of social norms perceives sanctions as cues for identifying expectations. However, defining sanctions in international law is problematic, though threats and coercive force are *prima facie* excluded[48].

Since customary norms evolve gradually through social practices and adapt to changing contexts, narratives help third parties understand the origins of a custom, how expectations are formed, and how they evolve, shaping normative expectations. A good narrator is sensible in plausibly recreating the characters' points of view to convey their personalities, goals, ways of thinking, and so on, to the audience. For a narrator, accurately recreating a character's perspective requires exploring the context of the States at a given historical moment, the values they uphold, and the assumptions they hold about each other. Through well-crafted narratives, we can gain insights into situations that would otherwise remain hidden.

Twining and Taruffo have championed the relevance of narratives or storytelling for evidential procedures, particularly as a tool of rational argumentation that allows structuring the materials provided for a case to make sense of them in a legally meaningful way (Twining, 2006, chap. 9). Narratives are also relevant for structuring a court's reasoning when assessing parties' actions and expectations in determining whether a customary practice exists, which involves a norm, and whether there was a defection by one of the parties. This comprises discerning and structuring the information collected to picture the context in which the customary practice developed, the peculiarities of that practice that may have led to expectations for the creation of a norm, and the attitudes displayed toward the practice. If, as Taruffo notes, good judicial narratives allow distinguishing facts from law (Taruffo, 2008, p. 194), in the case of customary norms, they serve to differentiate incidental actions from those tangled in the chain of concurrent actions that produce a custom and to individuate the peculiar attitudes that manifest normative attitudes and expectations[49].

Reconstructing the narrative of the practice does not mean producing a precise, descriptive history in which one can clearly distinguish when a convention arises between the parties and evolves into a proper social or customary norm[50]. Recon-

[48] With the exceptions listed in article 2.4 of the Charter of the United Nations and diverse Resolutions.

[49] As Sender and Wood expose, commentaries on conclusion 3 of the *Drafts* emphasize that courts must be careful in considering the context of States' actions, which includes those declarations proffered causally, in the heat of the moment, should not have the same weight as carefully considered declarations. Also, declarations by junior officials should carry less weight than those of senior officials (Sender & Wood, 2024, p. 85).

[50] Some scholars are preoccupied with the importance of distinguishing between the stage of norm formation and the stage of preservation, as if it were possible to identify the exact moment

structing the narrative, even the narrative of the practice of a social norm, involves something less ambitious.

The evidence collected by the courts and presented by the parties is hardly sufficient to reconstruct a precise and complete history of the practice[51]. Moreover, evidence doesn't speak for itself; it requires interpretation to understand metaphysical concepts such as customs, norms, and expectations. Therefore, a historical reconstruction is almost inevitably an interpretation of the facts, and if the authors of such a reconstruction are unaware of this, the result risks being highly biased. Instead, courts must construct hypotheses about the existence (or nonexistence) of a customary norm, assess and articulate the evidence, and try to assume the perspective of a member of the group in the ongoing history in which the practice occurs. This may be the closest to adopting the internal point of view and to a strategy that can help reduce the knowledge gap in social norms between a member of the group and an external observer. This perspective may permit picturing an act of self-interest coordination among group members and an act of egotistic self-interest that disregards others' interests and expectations[52]. That is, in the usual account of social norms,

---

when a customary norm emerges from a mere custom (Akehurst, 1975, pp. 4-5; Bobbio, 2010, pp. 60-61; Chiassoni, 2008, pp. 78-79). The distinction between the formation and the conservation or consolidation of a customary norm has also been a central argument in diverse paradigmatic cases, such as the *Lotus Case*, the *North Sea Continental Shelf Case*, the *Nuclear Tests Jurisdiction Case*, and the case of *Germany v. Italy about Jurisdictional Immunities*, among others. It has been stressed by Conclusion 1 of the *Draf* (although acknowledging its difficulties). However, insistence is unhelpful because, as Ullmann-Margalit emphasizes, norms don't arise at a specific time nor from a manageable number of acts. Instead, a better strategy is to verify whether the conditions for a norm's existence have been met (Ullmann-Margalit, 1977, p. 8). She takes this insight from Raz's indications about the existence of isolated norms. Joseph Raz (1990).

[51] This is not a form of holism that ranks evidential reliability below narrative coherence. Narratives serve as heuristic tools, not replacements for detailed evidential analysis. This does not mean endorsing a holistic approach to evidence over an atomistic one. I see them as complementary (*see* Accatino, 2014, pp. 45-52). The atomistic approach assesses the reliability of each piece of evidence, while the holistic approach evaluates a set of evidence related to a hypothesis (Ferrer Beltrán, 2007, pp. 45-46). The strength of a proof set depends on the reliability of its parts. Therefore, a narrative must be sound and grounded in reliable evidence to best support a hypothesis.

[52] The question of what motivates actors to follow social norms is rarely debated and often unclear, leading to cryptic accounts of why some agents break norms underdeveloped. The traditional economist's view, based on *homo œconomicus*, suggests that actors follow norms because the benefits of compliance outweigh those of defection, implying that defectors are mistaken or foolish. Behavioral economists note this view is short-sighted, as balancing costs isn't always the main motivator and is affected by cognitive biases. Furthermore, it tends to assume a decision-theoretic model of an isolated decision-maker. [Ullmann-Margalit, 1977, p. 7]. Instead, conventionalist accounts assume decision makers work within a strategic social-interaction framework, and avoiding the coordination solution can sometimes offer greater personal benefit. Consequently, Bicchieri claims social norms conflict with selfish motives and emerge from situations with conflicting interests and coordination incentives. Hence, she emphasizes that norms develop when a mixed-motivation game evolves into a coordination game in which actors prioritize coordination, recognizing that coordination is more rewarding in the long term because selfishness carries higher risks. Therefore, actors' interests should not be treated in a Manichean way. Actors do cherish coordination solutions and, most of the time, observe and enforce

the coordination solution that involves the customary norm has the agents tie their conception of self-interest not only to success but also to success within the group.

Observers of the norm not only internalize the norm as a valid norm of the group and follow it habitually and even somehow automatically[53], but they also regard following norms as a requisite to coordinate with others and be socially accepted, or at least not repudiated. Defectors usually continue to desire being accepted and repudiate sanctions; the difference from observers is that defectors have desires they cannot achieve by coordinating with others because, it tends to be the case, their desires imply losses for others. Hence, what characterizes the actions of non-cynical defectors[54] is that they make efforts to hide their actions, and thus, also benefit from others' perceptions of their cooperation. On the other hand, observers would be able to explain how the defector's actions fail to meet their expectations (descriptive and normative) and how that, in some way, implies losses on their behalf. Losses commonly involve the non-satisfaction of an economically quantifiable expectation, but this may not be obvious at first sight. Sometimes the losses are characterized as dignitary torts, as the law's doctrine calls them, such as damage to reputation, invasion of privacy, or breach of confidence. Considering customary norms involve interdependent actions that anchor conditional preferences, the losses can also be understood in terms of the efforts made by one of the parties, driven by justified expectations that the other party will and should act in a corresponding manner.

The best epistemologically justified hypothesis, the one that permits abductive inference of the probable existence of customary norms as the best explanation, will be the one comprising more reliable and diverse information supporting the existence of a recurrent, convergent practice, while also consistently justifying the existence of shared expectations, both descriptive and normative. It must allow for constructing a narrative that plausibly explains the factors that led to the formation of a norm within a group through repeated convergent practice, including the expectations created by the practice and the attitudes the parties demonstrate toward it[55].

---

social norms. But that does not mean they have abandoned their egoistic selves. If actors foresee a short-term gain from breaking a norm and believe they can get away with it, they will be tempted, and some will do it.

[53] Internalization of social or customary norms is a key in their continuity and in their seemingly natural functioning. That is, some norms, like the rules of grammar, are followed unconsciously, meaning that agents do not stop to evaluate the rule's validity each time they apply it. This idea underpins Wittgenstein's rule-following thesis. Additionally, as Guala emphasizes, the internalization of social norms through habit has been an important concept in social philosophy, and it partly involves agents' compromise with the norms (2016, p. 79). In this line, in a laboratory experiment, Guala concluded that when people produce a story of strategic joint interactions, this involuntarily produces a sort of internalized pressure for them, compelling them to abide by the equilibrium solutions reached and followed in the past (conventional social norm), even in the absence of external sanctions or possible negative consequences (Guala, 2013).

[54] A cynical defector is someone who openly breaks the norm and does little to hide it.

[55] The literature on social norms lacks specific methods for identifying actions that break norms and the reactions they elicit. Bicchieri proposes asking participants how they expect others to react to

Particularly for proving normative expectations, the main clues are displays of disapproval by other members of the group toward behavior of the sort committed by the alleged transgressor, and exhibitions of conduct whose particular manner of execution reasonably permits the inference of normative attitudes on the part of the actors involved in the practice. Undoubtedly, other hints can reinforce the latter: although not a requirement of normativity, longevity and consistency tend to be presumptive indicators that a pattern of recurrent behavior comprises more than mere courtesy or fads in the participants' thoughts. This can be relevant for distinguishing between cases involving a customary norm and those involving mere alleged *comitas gentium,* as argued in cases like *Portugal v. India*, the *Asylum case*, and other maritime law disputes. Also, as indicated, transgressors of norms tend to make efforts to hide their actions[56]. Likewise, it is not at all uncommon for some participants to emphasize that their actions are in accordance with a customary norm, and for other participants to heighten approval of those actions. This can also serve as evidence of the practice's normativity.

Consequently, a compelling reconstructive account of customary international norms requires an expansive understanding of sanctions[57] as indicators of normativity, a necessity reinforced by international law's ban on overt aggression and supported by philosophical analysis (e.g., Bicchieri, Brandom), demonstrating that even subtle manifestations of discontent or estrangement can serve as sanctioning cues for underlying normative practices.

6.3. Towards a Standard of Proof

The framework developed in the preceding sections clarifies what courts must prove — the existence of normative expectations constituted by state practice — and how to reason about such proof — through narrative reconstruction that integrates diverse evidence. Yet a crucial question remains: how much evidence suffices to justify a determination that a customary norm exists? The question is political rather than epistemological — though involving epistemic considerations[58]. Without a standard to prove a customary norm's existence, courts rely on subjective judgments of evidence rather than objective criteria.

Standards of proof are rules that specify the threshold of corroboration required for a justified conclusion. Social norm theory and epistemology provide conceptual

---

a hypothetical defection, as this better reflects social norms than asking whether they would condemn the action, which reflects personal values (2017, p. 72).

[56] Sometimes, this involves not so much concealing the breach as disguising it as compliance with the norm. This may involve merely resorting to rhetorical devices, but more sophisticatedly, doing something that formally or superficially appears to observe the norm while doing the opposite.

[57] Contrary to the expansive understanding of sanctions advocated, for instance, Kelsen reduced sanctions in international law to threats and acts of war from one State to another. Kelsen, 1956: 14, 22.

[58] Ferrer Beltrán, 2021, p. 24.

guidance on what counts as evidence and how to reason from it, but they do not specify decision thresholds. Inference to the best explanation, the reasoning method underlying narrative reconstruction, does not itself constitute a standard of proof. One narrative may be better than available alternatives yet still rest on sparse or low-quality evidence, yielding insufficient justification for a sound conclusion. International law faces this problem: the ICJ has not articulated explicit evidentiary standards, and the doctrine's emphasis on vague criteria such as «consistency of practice» offers little guidance. The inconsistency in the Court's reasoning across cases compounds the difficulty of extracting implicit standards from practice[59].

The need for articulated standards is particularly acute in customary international law. In national legal systems, well-established standards such as «preponderance of evidence» or «beyond a reasonable doubt», although vague, at least inform the decision-maker about the responsibility for what is at stake and acknowledge that there is a risk of error being distributed among the parties. The proof of customary international norms treats matters «pragmatically,» so to speak.

Although the two-element formula is often taken as a standard, this is not actually the case. As seen, it is more of a doctrinal and epistemic blueprint to guide their inquiries and interpretations. However, courts can be reckless in their treatment of this blueprint, neglecting one element altogether or giving it little weight compared to the others. Moreover, each element could be evaluated by its own standard. The *usus* can be quantified and evaluated against factors such as time and uniformity. The *opinio juris* demands greater conceptual and interpretive effort. Hence, the ICJ tends to establish, case by case, the threshold that evidence of a customary norm must meet, according to the stakes at issue. For instance, in the *North Sea Continental Shelf* case, the ICJ was rigorous in its examination of *usus*, holding that evidence of a customary norm must display extensive and virtually uniform State practice, accompanied by *opinio juris* of the parties. Meanwhile, in *Nicaragua v. United States*, while not articulating a standard, emphasis was placed on *opinio juris*, particularly on State declarations, while quietly acknowledging that *usus* was neither extensive nor virtually uniform.

As attested, the stakes vary dramatically: proving a customary norm governing diplomatic immunity differs from proving one regulating the use of force. The consequences of error — falsely «finding» a non-existent norm or failing to recognize an existing one — are not symmetric across domains. A standard appropriate for low-stakes ceremonial practices may be insufficient for norms whose violation could trigger armed conflict. And while the flexibility of assuming diverse thresholds can be regarded as pragmatically convenient, not knowing beforehand the sort of standard the ICJ will assume is contrary to the predictability of its rulings and to the notion of legal certainty (2021, p. 28).

---

[59] On standards of proof in general and their importance (*see* Ferrer Beltrán, 2021; Laudan, 2006). On the specific problem in international law (*see* Del Mar, 2012).

While a comprehensive treatment of evidentiary standards for customary international law exceeds this article's scope, the framework developed here offers a promising direction. The baseline standard is preponderance of the evidence: a customary norm should be deemed proven when, holistically assessed, the evidence makes its existence more probable than not[60]. However, this must be supplemented with qualitative criteria derived from the narrative reconstruction method. Specifically, the evidence must support a coherent narrative that explains both the convergence in State practice and the normative expectations underlying it. This narrative coherence test requires: (1) that the evidence includes a qualitative sample of a cross-section of States' practice by geography, legal systems, and level of development[61], and (2) must give particular weight to States whose interest or legal positions are most directly implicated by the norm (as the doctrine from the *North Sea Continental Shelf* has established), (3) that the narrative account for both conforming conduct and apparent deviations (explaining the latter as either violations subject to criticism or instances where the norm does not apply); and (4) that the evidence include manifestations of normative expectations — criticism, justification, enforcement responses — not merely behavioral regularity. This standard is more demanding than a simple count of practice instances, yet more determinate than vague appeals to «consistency.»

This proposal is preliminary and requires further development. First, the framework does not address how standards should vary with the stakes involved — whether norms governing the use of force require a higher threshold than those governing diplomatic protocol. While differentiated standards exist in other legal areas, setting thresholds for categories of customary international law requires normative analysis of error costs, not merely epistemological considerations. Second, the narrative coherence test, though more determinate than current doctrine, is still interpretive and won't resolve all disputes about the existence of a norm. Its value lies in offering courts a methodologically sound framework for reasoning, not in mechanically ensuring

[60] For present purposes, «preponderance of the evidence» means the existence of a customary norm is more probable than not (>0.5), considering only the two hypotheses advanced by the parties. *Inference to the best explanation* thus serves as a method for comparing these competing claims, not as a freestanding standard of proof.

[61] Conclusion 8 of the *Draft* demands that, for determining the existence of a customary norm, the practice must be general, in the sense of being sufficiently widespread, representative, and consistent. Goldsmith and Posner notice that the determination that 190 or so States engage broadly in a particular customary practice that supports a customary norm is practically impossible. As Sender and Wood emphasize, the requirement of generality must not be understood as a requirement of universality (2014, pp. 139-140). Hence, in practice, the determination of customary international law tends to be based on selective survey practice, which undeniably privileges the interests of major international powers (Goldsmith & Power, 2005, p. 24). The bias towards powerful States is reproachable in the practice of actual International Courts, but that does not mean that this representative sample cannot be done, considering random States to attest that a practice is widely held if the case is to attest the existence of a universal customary norm, like, for instance, the norm about the inviolability of diplomatic agents or the immunity *rationae personae*. The scope of the sample can be restricted when it is argued to be a regional or bilateral custom.

certainty, but at least in showing, in a more transparent way, the factual and causal reasoning courts use in determining the existence of a valid customary norm. These questions indicate that the standards of proof for customary international law remain an important area for future research, one in which epistemology, legal theory, and policy analysis must be integrated.

7. CONCLUSION

This article has argued that persistent confusion in customary international law reflects inadequate attention to the relationship between ontology and epistemology — conflating what constitutes customary norms with how to prove them. Customary norms are formed through shared empirical and normative expectations. Proving them requires abductive inference: reasoning from observable phenomena to the best explanation of the normative expectations that structure them.

The article has made three main contributions. First, it has reconceptualized the relationship between *usus* and *opinio juris* to show that these are not independent variables requiring separate proof, but components ascertained from the same procedure: practice provides evidence from which normative expectations are inferred. This resolves the chronological paradox of custom formation. Second, it clarifies what courts must identify when seeking *opinio juris*: not beliefs that conduct is already *law* (which would be circular), but normative expectations. Third, it has proposed narrative reconstruction as an appropriate method, coherently integrating State practice, statements, and response patterns to infer expectations.

The article has also outlined a preliminary approach to evidentiary standards. While comprehensive development requires further work, the framework developed suggests that the appropriate standard should combine a preponderance of the evidence and narrative coherence. Evidence must support a consistent explanation of both behavioral convergence and normative expectations across representative States. However, it must be emphasized that development of differentiated standards for diverse domains of customary international law — reflecting varying stakes and acceptable error costs — remains an important direction for future research.

While developed through analysis of customary international law, this framework has broader implications for legal theory. The challenge of proving norms that emerge from practice rather than authoritative enactment arises in any legal system, including common law, commercial practices (*lex mercatoria*), and indigenous legal orders, among others. Conceptual and methodological tools developed here apply wherever law depends on identifying shared normative expectations embedded in social practice.

BIBLIOGRAPHY

Accatino, Daniela (2014). «Atomismo y holismo en la justificación probatoria». *Isonomía*, *40*: 17-59.

Akehurst, M. (1975). «Custom as a Source of International Law». *British Yearbook of International Law, 47*(1), 1-53.

Alchourrón, Carlos & Bulygin, Eugenio (1981). «The Expressive Conception of Norms». Hilpinen, Risto (*ed*.). *New Essays in Deontic Logic*. Reidel Publishing Company.

Bannelier, Karine, Christakis, Theodore & Heathcote, Sarah (*eds.*) (2012), *The ICJ and the Evolution of International Law*. Routledge.

Bederman, David (2010). *Custom as a Source of Law*. Cambridge University Press.

Beham, Markus (2022). «Schrödinger's Custom». Merkouris, Panos, Kammerhofer, Jörg. & Arajärvi, Noora (*eds*.). *Theory, Practice, and Interpretation of Customary International Law*. Cambridge University Press.

Bentham, Jeremy (2010). *Of the Limits of the Penal Branch of Jurisprudence* (Philip Schofield *ed*.). Clarendon Press.

— (1977). *A Comment on the Commentaries and A Fragment of Government* (J. H. Burns & H. L. A. Hart *eds*.). Clarendon Press.

Bicchieri, Cristina (2006). *The Grammar of Society. The Nature and Dynamics of Social Norms.* Cambridge University Press.

— (2014). «Norms, Conventions and the Problem of Expectations». Catwright, Nancy. & Montuschi, Eleonora (*eds*.). *Philosophy of Social Science*. Oxford University Press.

— (2017). *Norms in the Wild*. Oxford University Press.

Bobbio, Norberto (2010). *La consuetudine come fatto normativo*. G. Giappichelli Editore.

Brady, Michael S. & Fricker, Miranda (*eds*.) (2016). *The Epistemic Life of Groups*. Oxford University Press.

Brandom, Robert (1994). *Making it Explicit*. Harvard University Press.

— (2000). *Articulating Reasons*. Harvard University Press.

Brennan, Geoffrey, Eriksson, Lina, Goodin, Robert, & Southwood, Nicholas (2013). *Explaining Norms*. Oxford University Press.

Bulygin, Eugenio (1991). «Algunas consideraciones sobre los sistemas jurídicos». *Doxa. Cuadernos de Filosofía del Derecho*, *9*: 257-279.

Canale, Damiano (2008). «Paradossi della consuetudine giuridica». Zorzetto, Silvia (*ed*.). *La consuetudine giuridica. Teoria, storia, ambiti disciplinari.* Edizioni ETS.

Catwright, Nancy & Montuschi, Eleonora (*eds*.) (2014). *Philosophy of Social Science*. Oxford University Press.

Celano, Bruno (2010). *Fatti istituzionali, consuetudini e convenzioni*. Aracne.

Chiassoni, Pierluigi (2008). «Tre buoni filosofi contro i cattivi costumi». Zorzetto, Silvia (*ed*.). *La consuetudine giuridica. Teoria, storia, ambiti disciplinari*. Edizioni ETS.

Chwe, Michael S. (2001). *Culture, Coordination, and Common Knowledge*. Princeton University Press.

Collins, Stephanie & Lawford-Smith, Holly (2016). «The Transfer of Duties. From Individuals to States and Back Again». Brady, Michael S. & Fricker, Miranda (*eds*.). *The Epistemic Life of Groups* Oxford University Press.

D'Amato, Anthony (1971). *The Concept of Custom in International Law*. Cornell University Press.

— (1987). «Trashing Customary International Law». *The American Journal of International Law*, *81*(1), 101-105.

d'Aspremont, Jean, (2011), *Formalism and the Sources of International Law*, Oxford University Press.

— (2022). «The Custom-Making Moment in Customary International Law». Merkouris, Panos, Kammerhofer, Jörg & Arajärvi, Noora (*eds*.). *Theory, Practice, and Interpretation of Customary International Law*. Cambridge University Press.

Del Mar, Kathleen (2012). «The International Court of Justice and standards of proof». Bannelier, Karine, Christakis, Theodore & Heathcote, Sarah (*eds.*). *The ICJ and the Evolution of International Law*. Routledge.
Ferrer Beltrán, Jordi (2005). *Prueba y verdad en el derecho*, 2nd ed. Marcial Pons.
— (2007). *La valoración racional de la prueba*, Marcial Pons.
— (2021). *Prueba sin convicción*. Marcial Pons.
García-Carpintero, Manuel (1996). *Las palabras, las ideas y las cosas*. Ariel.
Gardner, John (2012). *Law as a Leap of Faith*. Oxford University Press.
Goldman, Alvin (1999). *Knowledge in a Social World*. Clarendon Press.
González Lagier, Daniel (2005). *Quaestio facti*, Vol. I. Palestra Editores.
— (2022). *Quaestio facti*, Vol. II. Palestra Editores.
Guala, Francesco (2013). «The Normativity of Lewis Conventions», *Synthese*, 190, 3107-3122
— (2016). *Understanding Institutions*. Princeton University Press.
Guastini, Riccardo (2011). *Interpretare e argomentare*. Giuffrè Editore.
(2018). «Dos concepciones de normas», *Revus* [Online], (35). URL: http:// journals.openedition.org/revus/3810. https://doi.org/10.4000/revus.3810
Haack, Susan (2014). *Evidence Matters*. Cambridge University Press.
Hart, H.L.A. (1994). *The Concept of Law*, 2nd ed. Clarendon Press.
Hart, H.L.A. (2013). «Discretion». *Harvard Law Review*, 127(2), 652-665.
Hilpinen, Risto (ed.), 1981, *New Essays in Deontic Logic*, Reidel Publishing Company.
Johnson, Katie (2021). «The Nature and Context of Rules and the Identification of Customary International Law». *The European Journal of International Law*, *32*(4), 1167-1190.
Jones, Todd (2011). «Uncovering the "Cultural Meaning": Problems and Solutions». Steer, Daniel & Guala, Francesco (*eds.*). *The Philosophy of Social Science Reader*. Routledge.
Kelsen, Hans (1949). *General Theory of Law and State*, transl. of A. Wedberg. Harvard University Press.
Kelsen, Hans (1956). *Principles of International Law*. Rinehart & Company, Inc.
Kirgis, Frederic (1987). «Custom on a Sliding Scale». *The American Journal of International Law*, *81*(1), 146-151.
Kunz, Josef (1953). «The Nature of Customary International Law». *The American Journal of International Law*, 47(4), 662-669.
Lewis, David K. (2002). *Convention*. Blackwell Publishing.
Lipton, Peter (2000). «Inference to the Best Explanation». Newton-Smith, William H. (*ed.*). *A Companion to Philosophy of Science*. Blackwell Publishers.
Marmor, Andrei (2009). *Social Conventions. From Language to Law*. Princeton University Press.
Mattei-Gentili, Piero (2020a). «Social Facts & the Semantic Conception of Norms. Customary Norms as a Test of Ontology». *Phenomenology and Mind*, *19*, 242-254.
— (2020b). «The Quest for the *Opinio Juris*: An Analysis of Customary Law, From Hart's Social Rules to Expectations and Everything in the Middle». *Noesis*, *34*, 89-114.
— (2024). «Normas como significados. Una revisión al debate sobre ontología de normas y teoría jurídica». *Doxa. Cuadernos de Filosofía del Derecho*, *48*, 121-153.
— (2026). *La normatividad de la costumbre*. Marcial Pons.
Merkouris, Panos., Kammerhofer, Jörg. & Arajärvi, Noora (*eds.*) (2022). *Theory, Practice, and Interpretation of Customary International Law*. Cambridge University Press.
Moreso, José Juan, Navarro, Pablo & Redondo, María Cristina (2002). *Conocimiento jurídico y determinación normativa*. Palestra Editores.
Navarro, Pablo E. (2005). «Acerca de la inevitabilidad de la interpretación». *Isonomía*, *22*, 99-123.
Newton-Smith, William H. (ed.) (2000). *A Companion to Philosophy of Science*. Blackwell Publishers.
Perreau-Saussine, Amanda & Murphy, James B. (*eds.*) (2007). *The Nature of Customary Law*. Cambridge University Press.
Postema, Gerald (2007). «Custom in International Law: A Normative Account». Perreau-Saussine, Amanda & Murphy, James B. (*eds.*). *The Nature of Customary Law*. Cambridge University Press.

Postema, Gerald (2012). «Custom, Normative Practice, and the Law». *Duke Law Journal*, *62*, 707 - 738.
Raz, Joseph (1990). *The Concept of a Legal System*, 2nd ed. Clarendon Press.
Roberts, Anthea (2001). «Traditional and Modern Approaches to Customary International Law: A Reconciliation». *The American Journal of International Law*, *95*(4), 757-791.
Searle, John (1969). *Speech Acts*. Cambridge University Press.
Sender, Omri & Wood, Michael (2024). *Identification of Customary International Law*. Oxford University Press.
Shapiro, Scott (2006). «What is the Internal Point of View». *Fordham Law Review*, *75*(3), 1157-1170.
Steer, Daniel & Guala, Francesco (*eds*.) (2011). *The Philosophy of Social Science Reader*. Routledge.
Sunstein, Cass (2019). *How Change Happens*. The MIT Press.
Talmon, Stefan (2015). «Determining Customary International Law: The ICJ's Methodology between Induction, Deduction and Assertion». *The European Journal of International Law*, *26*(2): 417-443.
Tams, Christian (2015). «Meta-Custom and the Court: A Study in Judicial Law-Making», *The Law and Practice of International Courts and Tribunals*, *14*, 51-79.
Taruffo, Michele (2008). *La prueba*, transl. by L. Manríquez & J. Ferrer Beltrán. Marcial Pons.
Twining, William (2006). *Rethinking Evidence*, 2nd ed. Cambridge University Press.
Ullmann-Margalit, Edna (1977). *The Emergence of Norms*. Clarendon Press.
Vernengo, Roberto (1976). *Curso de teoría general del derecho*, 2nd ed. Cooperadora de Derecho y Ciencias Sociales.
Williamson, Timothy (2002). *Knowledge and Its Limits*, Oxford University Press.
Zorzetto, Silvia (ed.) (2008). *La consuetudine giuridica, Teoria, storia, ambiti disciplinari*. Edizioni ETS.

Quaestio facti. Revista Internacional sobre Razonamiento Probatorio
Quaestio facti. International Journal on Evidential Legal Reasoning
Sección: Ensayos
2026 I 11 pp. 123-140
Madrid, 2026
DOI: 10.33115/udg_bib/qf.i11.23275
Marcial Pons Ediciones Jurídicas y Sociales
© Nathalie Mousist
ISSN: 2604-6202
Recibido: 23/04/2026 | Aceptado: 22/06/2026 | Publicado *online*: 30/06/2026
Editado bajo licencia Reconocimiento 4.0 Internacional de Creative Commons

¿VEN LO MISMO EXPERTOS Y NOVATOS? SOBRE LA ESTRUCTURA DE LA EXPERIENCIA PERCEPTUAL Y LA FUERZA DE LAS INFERENCIAS

Nathalie Mousist
nmousist@uahurtado.cl
https://orcid.org/0009-0005-3715-1683

RESUMEN: Este trabajo sostiene que la diferencia entre el testimonio de novatos y el testimonio de expertos no se reduce *solo* a una cuestión de grados de conocimiento, sino que involucra una diferencia estructural en la experiencia perceptiva. A partir de la noción de penetrabilidad cognitiva, se argumenta que los expertos perciben de manera distinta debido a los estados cognitivos y la carga teórica que modulan la percepción. Esto configura el contenido mismo de la experiencia y posibilita el acceso a rasgos del entorno que permanecen inadvertidos para el observador no entrenado. Dicha diferencia tiene implicancias epistémicas relevantes: las creencias del experto se forman a partir de una experiencia perceptual más rica, pero también más opaca a terceros. Sobre esta base, se introduce una distinción entre la percepción como punto de partida del razonamiento y las cadenas inferenciales que se construyen a partir de ella, y se analiza dónde debe ubicarse el escrutinio epistémico y sus limitaciones.

* Este trabajo forma parte de mis compromisos en el Doctorado en Derecho de la Universidad Alberto Hurtado (UAH). Lo que aquí desarrollo fue motivado a partir del seminario "Kant, Conceptos y Experiencia: una mirada contemporánea", dictado en el Doctorado de Filosofía (UAH), dirigido por Eduardo Molina y Francisco Pereira. Las discusiones allí sostenidas, junto a mis compañeros y compañeras, abrieron una línea de reflexión que resultó decisiva para lo que aquí se aborda, a todos ellos mis agradecimientos. Una versión preliminar de este manuscrito fue presentada en las Jornadas Nacionales de Filosofía del Derecho (Arica, 2025), allí agradezco los agudos comentarios y observaciones de Daniela Accatino, Flavia Carbonell y Marcela Chahuan, que contribuyeron significativamente a mejorar sus versiones posteriores. Finalmente, a quienes revisaron anónimamente ese trabajo.

PALABRAS CLAVE: percepción experta; penetrabilidad cognitiva; justificación epistémica; razonamiento probatorio; prueba pericial.

DO THEY SEE THE SAME THING? THE PERCEPTUAL EXPERIENCE OF NOVICES AND EXPERTS

ABSTRACT: This paper argues that the difference between the testimony of novices and that of experts is not merely a matter of degrees of knowledge, but rather involves a structural difference in perceptual experience. Drawing on the notion of cognitive penetrability, it is argued that experts perceive differently due to the cognitive states and theoretical framework that modulate perception. This shapes the very content of experience and enables access to features of the environment that remain unnoticed by the untrained observer. This difference has significant epistemic implications: the expert's beliefs are formed from a perceptual experience that is richer, but also more opaque to third parties. On this basis, a distinction is introduced between perception as the starting point of reasoning and the inferential chains constructed from it, and an analysis is conducted of where epistemic scrutiny should be directed and its limitations.

KEYWORDS: expert perception; cognitive insight; epistemic justification; evidentiary reasoning; expert testimony.

SUMARIO: 1. INTRODUCCIÓN.— 2. LA MISMA ESCENA, DISTINTAS PERCEPCIONES.— 3. IMPLICANCIAS EPISTEMOLÓGICAS DE LA PENETRABILIDAD COGNITIVA EN LA EXPERIENCIA PERCEPTIVA.— 4. BENEFICIOS Y RIESGOS DE LA PERICIA PERCEPTUAL.— 5. CADENAS INFERENCIALES Y JUSTIFICACIÓN PERCEPTUAL.— 6. CONCLUSIONES.— BIBLIOGRAFÍA.

1. INTRODUCCIÓN

En la epistemología general se considera, casi sin tapujos, que la diferencia entre el testimonio de novatos[1] y el testimonio de los expertos se reduce sólo a una cuestión de grados de conocimiento. Así, por una parte, se considera que habría una «ventaja posicional» del testigo por haber estado en el lugar y en el momento preciso para percibir algo relevante acerca de lo que informa; por la otra, se asume que habría una «ventaja disposicional» del experto que posee y está en disposición de utilizar adecuadamente un conocimiento especializado, una *expertise* relevante (Williams, 2002, p. 42). Aquello obedece a que los expertos saben más que los novatos sobre determinadas áreas del saber científicas o técnicas. En esta línea, se sostiene que de lo que se trata es de una «ventaja epistémica» del experto sobre la audiencia (Gelfert, 2014, p. 20).

[1] Uso la expresión «novato» en el sentido de no-experto.

Estas posiciones, que omiten diferencias epistémicas cualitativas, han sido acogidas en la filosofía del derecho y la teoría de la prueba (Accatino, 2018, p. 10). Una de las razones quizás más evidente de esta recepción es que tales posturas ofrecen un marco explicativo de carácter general para comprender las dificultades en la valoración del conocimiento experto en el ámbito judicial[2], en la medida en que se lo concibe como un tipo de conocimiento que, por su contenido y forma, excede las competencias cognitivas ordinarias de los tribunales.

Comparto parcialmente esta perspectiva en la medida en que reconoce una asimetría en la cantidad de conocimiento, información disponible y dominio técnico. Sin embargo, investigaciones contemporáneas provenientes de otros marcos teóricos —especialmente en filosofía de la percepción y ciencias cognitivas— advierten que esta caracterización resulta insuficiente para dar cuenta del fenómeno en su complejidad. Estos estudios sugieren que la experticia no se limita *solo* a una acumulación cuantitativa de datos, sino que introduce una modificación estructural en la forma en que los sujetos perciben el mundo. Dicho de otro modo: los expertos no solo saben más, sino que *ven distinto* (Dror, McCormack, y Epstein, 2015).

La aproximación desarrollada en este trabajo presupone una determinada concepción acerca de la relación entre epistemología y ciencias empíricas. En particular, se asume que los hallazgos de disciplinas como la filosofía de la percepción y ciencias cognitivas pueden contribuir a la comprensión de problemas epistemológicos relativos al conocimiento experto sin que ello implique reducir las cuestiones normativas a hechos empíricos. En este sentido, el enfoque se aproxima a lo que Haack (1997) denomina un naturalismo reformista *a posterior*, según el cual la epistemología conserva su autonomía normativa, aunque debe mantenerse informada por los resultados de las ciencias empíricas. Asimismo, presupone una cierta continuidad entre los procesos cognitivos que operan en la ciencia y aquellos que intervienen en otros contextos de producción y evaluación del conocimiento, incluido el ámbito jurídico.

El objetivo de este trabajo es indagar si, además de las diferencias de grado, existen diferencias cualitativas y estructurales en la manera en que novatos y expertos perciben el mundo[3]. Sostengo que esta diferencia no se agota en la asimetría cuantitativa de conocimiento proposicional, sino que también involucra una modificación en la estructura de la experiencia perceptiva. Esta hipótesis me permite explorar cómo

[2] Estas dificultades han sido ampliamente desarrolladas en la literatura sobre valoración de la prueba pericial. Suelen vincularse, entre otros factores, a la dependencia cognitiva de la jueza hacia el perito, la asimetría epistémica entre ambos, la autoridad epistémica que se atribuye al experto, el déficit de control respecto de los métodos utilizados y las limitaciones para evaluar la fiabilidad del testimonio experto. Véase Vázquez (2015; 2020), Gascón Abellán (2013; 2019), Haack (2014; 2020), Herdy (2020), entre otros.

[3] La respuesta a esta pregunta tiene impacto en la valoración de la prueba pericial, en tanto la manera en que los expertos y novatos procesan la información del mundo condiciona no solo la calidad epistémica de sus testimonios, sino también podría profundizar la asimetría epistémica afectando así la justificación de las inferencias probatorias. Sobre esto volveré en la sección V de este trabajo.

la experiencia del experto incide en la formación de sus creencias, en la configuración de sus cadenas inferenciales y en el tipo de justificación perceptiva que están en condiciones de ofrecer (Jenkin y Siegel, 2015).

El análisis de este problema, por una parte, refuerza la idea de que la percepción no es un simple mecanismo receptivo o pasivo de registro de estímulos, sino un proceso cognitivamente penetrado, cuyo contenido representacional se encuentra modelado por esquemas cognitivos, hábitos interpretativos, habilidades conceptuales y una historia de entrenamiento que orienta selectivamente la atención y configura lo que cuenta como perceptible[4]. Por otra parte, permite resaltar que esta estructuración diferencial de la experiencia da lugar a asimetrías epistémicas sustantivas entre el testimonio de novatos y el de expertos[5], que no solo afectan la fiabilidad de sus juicios, sino también las condiciones bajo las cuales sus creencias pueden considerarse justificadas y la fuerza epistémica de las inferencias que derivan de su experiencia.

Para abordar el interrogante, primero (2), presento una escena donde un mismo hecho es observado por sujetos con distintos niveles de *expertise*, con el fin de mostrar cómo la experiencia perceptiva puede variar frente a un mismo estímulo. En segundo lugar (3), sitúo la discusión en el marco de la filosofía de la percepción, atendiendo especialmente a la noción de penetrabilidad cognitiva y su papel en la configuración de la experiencia perceptual experta. En tercer lugar (4), analizo los beneficios y contra de la pericia perceptual y sostengo que tiene un carácter ambivalente, en tanto constituye simultáneamente una condición de posibilidad de su valor epistémico y una fuente potencial de sesgos. Finalmente, (5), examino las implicancias de este enfoque para la teoría de la prueba, en particular en relación con el alcance del control judicial y la valoración de la prueba pericial.

2. LA MISMA ESCENA, DISTINTAS PERCEPCIONES[6]

Los primeros en llegar fueron algunos vecinos, alertados por la puerta entreabierta y un olor extraño que salía del interior del departamento. En el centro del dormitorio se encontraba el cuerpo de un hombre. Los vecinos se detuvieron a unos pocos metros y vieron lo que cualquier observador no entrenado percibiría de inmediato: un cuerpo inmóvil, la piel ligeramente amoratada y el desorden de una escena inquie-

[4] Como se explicará más adelante, esta idea cuya formulación clásica se remonta a Kant, fue posteriormente reelaborada en la filosofía de la ciencia del siglo xx (Hanson, 1977) y retomada en la literatura contemporánea (Siegel, 2015, entre otros), en relación con la idea de la «carga teórica» de la percepción.

[5] Hablamos en sentido epistémico de testimonio cuando se obtiene conocimiento de lo que otros dicen, siempre que un agente expresa, mediante un acto de comunicación, información a otro (la audiencia), que forma sobre esa base sus creencias (Vázquez, 2015, p. 50 y ss.).

[6] La escena ilustrativa que se propone es la adaptación de un ejemplo ofrecido por Wahlberg y Dahlman (2021, p. 54).

tante. Comentaron si la muerte habría ocurrido «hace mucho», si pudo tratarse de «una caída», «un ataque al corazón» o un «robo que salió mal». Pero más allá de estas conjeturas, nada adicional se volvió perceptible ante sus ojos, solo había un cuerpo rígido con los ojos abiertos.

Minutos más tarde llegaron dos médicas forenses. La escena era la misma: idéntico cuerpo, idéntica iluminación, idénticas condiciones ambientales, nada había cambiado. Sin embargo, la experiencia perceptual que tuvieron fue diferente a la de los vecinos. Una de las forenses ajustó la lámpara portátil y dirigió la luz hacia los ojos de la víctima. Inmediatamente advirtió un conjunto de petequias (manchas rojas) distribuidas en la conjuntiva ocular. Las manchas, invisibles o irrelevantes para los vecinos, se presentaban ante las expertas como un patrón familiar: un indicio recurrente asociado a la muerte por asfixia mecánica.

La presencia de petequias había sido inadvertida para los vecinos, y si acaso alguno las percibió, no tendrían significado alguno[7]. Sin embargo, ello junto con la coloración del cuello y la ausencia de lesiones compatibles con una caída, permitió formular una primera hipótesis altamente informada: muerte por estrangulamiento. Las expertas no solo informaron *más* de lo que podría informar un vecino, sino que *observaron otra cosa* porque su experiencia perceptual estaba organizada por años de entrenamiento que orientaron selectivamente la atención hacia rasgos significativos del estímulo.

La escena busca ilustrar, de manera sencilla, la tesis de que ver no es simplemente mirar, sino percibir selectivamente en virtud de estados cognitivos y habilidades adquiridas. A continuación, desarrollaré esta idea a partir de la noción de penetrabilidad cognitiva, según la cual los estados cognitivos pueden influir de manera sistemática en qué rasgos del entorno se vuelven perceptualmente salientes. En esta clave la diferencia entre novatos y expertos se entiende no solo como asunto entre mayor o menor conocimiento, sino como una divergencia estructural en la fenomenología perceptiva y en la configuración de sus contenidos representacionales.

3. IMPLICANCIAS EPISTEMOLÓGICAS DE LA PENETRABILIDAD COGNITIVA EN LA EXPERIENCIA PERCEPTIVA

La percepción no se agota en lo observado, sino que también involucra las características propias del sujeto que observa. En términos generales, observar implica

[7] Incluso si la jueza hubiera concurrido al lugar del hecho, por ejemplo, mediante una inspección ocular, es altamente improbable que hubiera advertido las petequias, y aun en caso de percibirlas no podría atribuirle significado probatorio sin la mediación del perito. La identificación y la interpretación de este tipo de indicadores (en este caso, fisiológico) requiere esquemas conceptuales y hábitos perceptivos adquiridos exclusivamente mediante formación especializada y experiencia reiterada en autopsias y análisis forenses. Esta limitación no es contingente, sino estructurada y explica por qué la inspección ocular judicial no elimina la necesidad de convocar a un experto.

la capacidad de captar a través de los sentidos, ciertos rasgos de una situación que ocurre en el mundo. En este sentido, cuando evaluamos una percepción, solemos preguntarnos si el sujeto estuvo en condiciones adecuadas para percibir lo que dice haber visto, oído o sentido y cómo ello repercute en la fuerza de su testimonio. Así, consideramos variables como la distancia, la iluminación o si sus capacidades sensoriales permiten un acceso fiable a la situación. Este tipo de análisis presupone que la calidad de la percepción depende de las condiciones externas bajo las cuales se produce la observación.

Sin embargo, nuestra percepción del mundo no depende únicamente de los estímulos que recibimos a través de nuestros sentidos. En filosofía de la ciencia, este problema ha sido abordado bajo la noción de «carga teórica de la observación», que destaca que lo percibido está mediado por conceptos, expectativas y formas de organización cognitiva. Esta idea tiene su antecedente en Kant, quien sostuvo que la experiencia no es un mero reflejo del mundo, sino el resultado de la cooperación entre la sensibilidad y el entendimiento: los datos sensibles solo se vuelven inteligibles bajo formas y conceptos que el sujeto aporta[8].

En la filosofía contemporánea, esta línea fue desarrollada, entre otros, por Sellars (1956) y McDowell (1994), a partir de la crítica al «Mito de lo Dado», según la cual no existen contenidos perceptuales puros o independientes de las capacidades conceptuales del sujeto[9]. La experiencia perceptiva esta ya «impregnada» de conceptualidad y solo puede desempeñar un rol justificativo en la medida en que se inscribe en el espacio de las razones, y no como un mero impacto causal en los sentidos. Bajo este enfoque, la percepción deja de entenderse como un *input* neutro y pasa a concebirse como un proceso estructurado por las capacidades conceptuales y antecedentes cognitivos del agente.

En este marco, la noción de penetrabilidad cognitiva puede entenderse como una reformulación contemporánea de este problema, en la medida en que permite precisar los modos en que los estados cognitivos influyen en la experiencia perceptual. Más específicamente, la penetrabilidad cognitiva designa un tipo de influencia causal que ciertos estados cognitivos[10] ejercen sobre la experiencia visual. Entonces, una experiencia visual es cognitivamente penetrable cuando, es nomológicamente posible que dos sujetos (o un mismo sujeto en diferentes circunstancias o momentos) tengan experien-

[8] En este trabajo por «esquemas o habilidades conceptuales» se entiende el conjunto de herramientas conceptuales (conceptos, clasificación y criterios) que los sujetos emplean para identificar, organizar y describir lo que perciben. Estas llamadas «categorías kantianas» constituyen el aporte imprescindible de la razón a la experiencia (Torretti, 2012).

[9] Sobre este punto existe un amplio debate entre posiciones conceptualistas y no conceptualistas acerca de la naturaleza del contenido perceptual (véase, por ejemplo, Toribio, 2007), con implicancias relevantes para la explicación de la percepción de sujetos sin capacidades lingüísticas, como niños y niñas pre-lingüísticos y animales no humanos.

[10] Los estados cognitivos penetrantes pueden ser las emociones, los deseos, las aversiones, las expectativas. Me centro aquí en estados cognitivos doxasticos como las creencias.

cias con contenidos diferentes frente a los mismos estímulos distales[11] y bajo las mismas condiciones externas, debido exclusivamente a diferencias en sus estados cognitivos previos. En estos casos, los estados «penetrantes» modulan qué rasgos del estímulo distal son captados, orientando la percepción hacia determinados aspectos del entorno.

En contextos de experticia, esta modulación adquiere un carácter selectivo. Aunque el estímulo distal permanezca invariable para el experto y el novato, el trasfondo cognitivo del experto hace que él dirija su atención hacia rasgos del entorno que el novato, simplemente, no registra. Las variables que explican esta selectividad son dos. Por un lado, la fijación visual moldeada por la experiencia y, por el otro, la atención encubierta. Pues se trata de mecanismo mediante los cuales la pericia determina no solo *qué* se ve, sino también *cómo* se ve.

En el caso del estrangulamiento esto se puede explicar del siguiente modo. Los vecinos y las médicas forenses observan el mismo cuerpo bajo las mismas condiciones, pero mientras los vecinos perciben rasgos globales (color de la piel, cuerpo quieto, desorden), las médicas fijan su atención en detalles más finos y precisos: la distribución de petequias en los glóbulos oculares y la coloración del cuello. Para ellas, estos rasgos son inmediatamente salientes y significativos, no porque tengan un acceso privilegiado al objeto, sino porque sus estados cognitivos previos (patrones aprendidos, categorías técnicas, hábitos atencionales) modulan qué aspectos del estímulo distal capturar y con qué organización conceptual aparecen. Así, aunque todos miran la misma escena, el contenido de la experiencia visual difiere. Esta modulación explica, incluso, por qué distintos expertos pueden estructurar de manera diversa una misma experiencia perceptiva. En efecto, las diferencias en los marcos conceptuales, las metodologías adoptadas o los criterios relativos al riesgo de error inciden en qué aspectos del fenómeno son detectados como relevantes y en cómo son interpretados. Estas divergencias, no necesariamente reflejan desacuerdos epistémicos en sentido estricto, sino variaciones legítimas dentro de la práctica científica en la forma de identificar y caracterizar los fenómenos (Páez, 2026, p. 50).

En este contexto, la «pericia perceptual» se refiere a una forma de experiencia que implica la percepción sensorial y se desarrolla a través del aprendizaje y la práctica, lo que permite a los expertos sobresalir en la percepción dentro de su dominio especializado. Los expertos perceptuales no solo son más rápidos y precisos en sus

[11] En filosofía de la percepción se denomina «estímulos distales» a los objetos o propiedades del mundo externo que originan (a distancia) nuestras experiencias perceptivas. Se distinguen de los «estímulos proximales», que son los patrones de estimulación que alcanzan directamente a los órganos sensoriales (por ejemplo, la distribución de luz en la retina). Mientras el estímulo distal corresponde al estado de cosas objetivo (como un árbol, una taza de café en la mesa o un conjunto de manchas rojas en el glóbulo ocular), el estímulo proximal es la base fisiológica inmediata de la percepción. Esta distinción resulta relevante en el análisis de la pericia perceptual, pues permite mostrar cómo dos sujetos pueden estar expuestos al mismo estimulo distal y, sin embargo, tener experiencias visuales diferentes. Tales diferencias no se explican por variaciones en el mundo, sino por modulaciones internas vinculadas a la atención, esquemas conceptuales, entrenamiento previo o la estructura cognitiva del perceptor. Véase Burge (2010), Siegel (2010).

observaciones, sino que también poseen una mayor sensibilidad a los detalles finos. Ven las cosas más rápidamente, distinguen patrones que otros no pueden detectar y hacen comparaciones rápidas entre características perceptibles que para otros son difíciles de entender (Ransom, 2020). En estos casos, la experiencia perceptual es simultáneamente sensorial y profundamente cognitiva, está cargada de teoría y conocimiento previo, lo que aporta un valor epistémico que conduce a una comprensión más profunda del fenómeno observado (Stokes, 2021).

4. BENEFICIOS Y RIESGOS DE LA PERICIA PERCEPTUAL

La discusión sobre la penetrabilidad cognitiva es relevante en el contexto de la experticia perceptual. En muchos casos —y de manera paradigmática en el ámbito pericial— la influencia de estados cognitivos previos sobre la experiencia visual no solo es un defecto epistémico, sino que también parece ser la condición de posibilidad de que una percepción sea epistémicamente valiosa. Por ello, el problema central no es si la experiencia experta está cognitivamente penetrada, sino qué consecuencias epistémicas se siguen de esa penetración.

Desde una perspectiva favorable, la penetrabilidad cognitiva permite comprender por qué los expertos perciben rasgos del entorno que permanecen inaccesibles para los observadores no entrenados. El ejemplo clásico es el de la radióloga que detecta en una radiografía un tumor o una mancha que un observador lego no logra advertir[12]. La diferencia entre ambos radica principalmente en la capacidad para organizar perceptualmente la información disponible a la luz de conocimientos previos, patrones aprendidos y hábitos atencionales específicos. Gracias a esta modulación cognitiva, la experiencia perceptual del experto porta un contenido epistémicamente más rico y relevante.

Sin embargo, esta estructura que explica el valor epistémico de la percepción experta introduce un riesgo significativo. Algunos autores han señalado que la penetrabilidad cognitiva puede también deteriorar la justificación perceptual, al generar formas de circularidad entre creencia y experiencia. En términos generales, cuando una experiencia es penetrada cognitivamente, el mundo se presenta como si fuera de cierta manera solo porque esa es la forma en que la creencia penetrante presenta el mundo (Siegel, 2012). En estos casos, algunos filósofos[13] sostienen que la experiencia reduce su capacidad de proporcionar una justificación para creer en su contenido,

[12] Es un ejemplo utilizado por Siegel (2009).

[13] McDowell (1994) y Sosa (2007) forman parte de una corriente que sostiene que la experiencia perceptiva no puede ser el resultado de una inferencia. En el caso de McDowell, antes de su giro posterior (2009), esta posición se desarrolla en *Mind and World*, donde defiende que la experiencia perceptiva ya se encuentra en el «espacio de las razones». En este marco, ver que *p* es el caso implica estar en una posición que justifica creer que *p*, de modo tal que la experiencia no es una conclusión inferencial, sino una fuente inmediata de justificación. Para una discusión más profunda sobre el alcance

fenómeno conocido como «degradación epistémica». Una experiencia epistémicamente degradada es aquella cuya capacidad justificatoria, incluso *prima facie*, se ve disminuida debido al modo en que se forma (Chudnoff, 2019).

Este problema ha sido formulado con claridad por Susanna Siegel (2009; 2017), quien sostiene que la capacidad justificatoria de una experiencia perceptual depende, al menos en parte, de su etiología, esto es, de los procesos y condiciones que la originan. Cuando la experiencia tiene un «pasado accidentado», esto es, cuando resulta de procesos cognitivos que serían epistémicamente defectuosos si produjeran una creencia en lugar de una experiencia, su fuerza justificatoria se ve degradada. En tales casos, la experiencia no cumple su función epistémica habitual de poner a prueba las creencias del sujeto frente al mundo, sino que introduce una estructura circular de confirmación.

Para ilustrar este punto Siegel propone, entre otros, el siguiente caso. Si (i) Jill cree que Jack está enojado y, a partir de esa creencia, (ii) tiene una experiencia visual de ira en el rostro de Jack, entonces (iii) Jill confirma su creencia inicial. Se trata de una situación epistémicamente perniciosa, ya que la experiencia deja de funcionar como un tribunal independiente: en lugar de contrastar la creencia con el mundo, la refleja. En estos casos, los estados penetrantes manipulan la experiencia a su favor, de modo que el sujeto no introduce ninguna justificación nueva o independiente. La experiencia aparenta portar evidencia adicional, pero en realidad solo confirma las expectativas que la generaron.

Según Siegel, en este tipo de situaciones existe una base aparente de apoyo epistémico para creer que Jack está enojado, asociada a una experiencia visual. Sin embargo, dicha experiencia está causada por una premisa infundada —por ejemplo, miedo o una sospecha infundada—, lo que hace que el respaldo epistémico que ofrece a la creencia resultante caiga por debajo de una línea de referencia adecuada. Este es, precisamente, el mecanismo de degradación epistémica.

En el ámbito de la experticia, esta forma de degradación suele manifestarse como un sesgo de confirmación, con consecuencias directas para la fiabilidad del testimonio experto. Aunque suele suponerse que los expertos son particularmente imparciales y objetivos, diversos estudios muestran que también son vulnerables a sesgos que nacen, precisamente, de los mecanismos cognitivos que sostienen su pericia. En este sentido, Itiel Dror (2020, p. 7999) [14] señala que

---

y los límites de la tesis según la cual «ver es creer» así como los desarrollos posteriores de McDowell, véase Pereira (2021).

[14] Para profundizar véase Dror (2020). En este trabajo el autor identifica seis falacias sobre la naturaleza de los sesgos: es una cuestión ética, solo aplica a manzanas podridas, los expertos son imparciales e inmunes, la tecnología elimina el sesgo, el punto ciego y la ilusión del control. A continuación, discute y conceptualiza ocho fuentes de sesgos en tres categorías: a) factores relacionados con el caso específico y análisis, que incluyen datos, los materiales de referencia y la información contextual, b) factores relaciones con la persona especifica que realiza el análisis, que incluye las tasas de base de a experiencia previa, los factores organizacionales, la educación y la capacitación, y los factores personales, c) la arquitectura cognitiva y la naturaleza humana que nos impacta a todos.

> La experiencia y formación hacen que los expertos presenten una atención más selectiva, utilicen fragmentación y esquemas (actividades típicas y su secuencia) y se basen en heurísticas y expectativas que surgen de experiencias base pasadas, utilizando toda una gama de procesos cognitivos de arriba hacia abajo que crean suposiciones y expectativas *a priori*.

Esto explica cómo el contexto y la cognición dan forma silenciosamente al juicio de los expertos. Cuando un científico forense, por ejemplo, sabe quién es el sospechoso o qué espera la policía (esto es una fuente de sesgo relacionada con la información contextual), su interpretación de los mismos datos puede cambiar, generando una cascadas de sesgos o efectos de «bola de nieve» a lo largo de una investigación[15]. No obstante, tal como destacan Dror (2020) y Dahlman y Walbergh (2021), estos sesgos no son el resultado de decisiones intencionales, sino que operan de manera inconsciente como parte del funcionamiento del cerebro humano: la atención, la memoria y la inferencia están guiadas por la experiencia, expectativas e incluso por presiones organizacionales. En este sentido, los expertos no solo «ven» la evidencia, sino que la construyen a través de sus filtros cognitivos.

Estas limitaciones justifican la necesidad de salvaguardas institucionales, como el cegamiento, la secuenciación de información o la revisión independiente, destinadas a minimizar el impacto de sesgos inevitables[16]. Sin embargo, reconocer estos riesgos no equivale a concluir que la penetrabilidad cognitiva sea, en sí misma, un defecto epistémico. Por el contrario, sugiere una ambivalencia en la percepción experta: la misma permeabilidad de la experiencia a estados cognitivos previos que permite detectar patrones relevantes es la que abre la posibilidad de sesgos y distorsiones.

Sería, por ello, un error concluir que la penetración cognitiva debe ser eliminada. Existen buenas razones para pensar que la permeabilidad de la percepción al trasfondo cognitivo es precisamente lo que hace valiosa la experiencia del experto. Su percepción es más rica y sustantiva *porque* está modulada por su conocimiento previo. Incluso, la resistencia a modificar ciertas intuiciones frente a evidencia contradictoria —a menudo interpretada como un sesgo— puede constituir una forma de tenacidad epistémica, esto es, un rasgo virtuoso cuando deriva de patrones perceptuales finamente entrenados y no de prejuicios infundados.

Si esto es así, se puede afirmar que la penetrabilidad cognitiva de los expertos en las experiencias perceptivas tiene una doble cara: puede generar sesgos, pero es

---

[15] Un caso paradigmático es de Brandon Mayfield. Tras los atentados de Madrid de 2004, el FBI identificó erróneamente a Mayfield como autor a partir de una coincidencia dactilar considerada "100%" por varios peritos. Sin embargo, la Policía Nacional Española había determinado que la huella no correspondía a él. A pesar de esta información, los examinadores del FBI insistieron en su identificación inicial. Investigaciones posteriores atribuyeron el error, entre otras causas, a un sesgo de confirmación y a una dinámica de circularidad interpretativa, en la que los peritos ajustaron a la interpretación de la huella para hacerla coincidir con el sospechoso inicialmente seleccionado, ignorando discrepancias relevantes y validando sus conclusiones mediante procesos de verificación no independientes.

[16] Para profundizar en problemas de sesgo cognitivo, sesgo contextual y sesgo del observador y formas prácticas de mitigar estos problemas, véase Dror, McCormack y Epstein (2015).

también la condición misma de su capacidad para detectar rasgos que los novatos no ven. Reconocer esta ambivalencia permite superar una visión, quizás, simplista según la cual toda influencia cognitiva es un defecto, y abre la puerta a una concepción más matizada en la que la pericia perceptual, adecuadamente regulada, funciona como un recurso epistémico indispensable, y no como una amenaza[17].

Esto permite, a su vez, reformular el problema en un plano distinto. Si la penetrabilidad cognitiva es una condición estructural del poder epistémico del experto, y no un defecto eliminable, entonces, —más allá de la cuestión relativa de los sesgos y la necesidad de mecanismos orientados a mitigarlos— me interesa aquí examinar en qué medida la experiencia perceptual puede ser evaluada, bajo qué condiciones y sus eventuales limitaciones.

En lo que sigue, sostendré que la clave para abordar esta cuestión, es analizar primero el lugar que ocupa la experiencia perceptual en relación con las cadenas inferenciales que el experto articula a partir de ella, ¿cómo pasa de lo que ve a lo que afirma? ¿dónde se ubica el control? Esto me permitirá mostrar que las dificultades asociadas a la percepción experta no se agotan en ella ni desaparecen a nivel inferencial. Sin embargo, las formas en que dicha experiencia se articula y comunica en el razonamiento probatorio ofrecen mayores posibilidades de explicitación y, con ello, de escrutinio epistémico.

5. CADENAS INFERENCIALES Y JUSTIFICACIÓN PERCEPTUAL

En el razonamiento probatorio, la valoración de la prueba se estructura a partir de inferencias que permiten a las juezas arribar a conclusiones epistémicamente fiables. Como señala González Lagier, esta mediación —denominada enlace o conexión— se despliega entre los elementos del juicio y la hipótesis a probar, y puede adoptar un carácter empírico, normativo o conceptual (González Lagier, 2022, pp. 16-17). En el caso de las inferencias empíricas o inferencias probatorias epistémicas, el enlace consiste en una generalización que correlaciona hechos como los descriptos en las premisas con los hechos afirmados en la conclusión, apoyándose en asociaciones observadas con anterioridad, ya sea a través de máximas de la experiencia o de conocimiento científicos (González Lagier, 2018, p. 23). De este modo, la solidez de una cadena inferencial depende, en gran parte, de la calidad de los enlaces que la componen: la adecuación de las generalizaciones empleadas, la coherencia del razonamiento, la consideración de hipótesis alternativas y la sensibilidad a la contra evidencia.

[17] Es indudable que los sesgos constituyen una fuente de error, cuyos efectos negativos han sido ampliamente documentados en el ámbito jurídico. Sin embargo, pese a la evidencia disponible, los sistemas judiciales no han incorporado de manera sistemática mecanismos adecuados para abordarlos, siendo en muchos casos tolerados o ignorados. Para un análisis reciente, véase Páez (2026). En este trabajo, no obstante, el foco no está puesto en los sesgos como déficit, sino en el valor epistémico de la penetrabilidad cognitiva como condición de posibilidad de la pericia perceptual.

En el caso de las inferencias expertas la calidad de esos enlaces depende, además de la acumulación de conocimiento proposicional, del modo en que la experiencia perceptual los alimenta, proporcionando patrones, relaciones y rasgos que permiten construir conexiones más finas y precisas. En este sentido, el experto no solo está en condiciones de inferir más, sino mejor, en la medida en que sus inferencias se apoyan en generalizaciones mejor calibradas. Esto presupone una relación de continuidad entre percepción y cognición en la medida en que, como se viene argumentando, la experiencia perceptual experta se encuentra cognitivamente modulada y proporciona, por ello, una base epistémicamente más rica para la construcción de inferencias.

Esta continuidad es compatible con una distinción conceptual entre la experiencia perceptual y las cadenas inferenciales que se articulan a partir de ella. Mientras que la primera se refiere al modo en que ciertos aspectos del entorno se vuelven perceptualmente salientes o accesibles a los expertos; las segundas, consisten en las conexiones que permiten relacionar esos datos con una hipótesis y justificar una conclusión. Desde esta perspectiva, la tesis de la penetrabilidad cognitiva se refiere a los factores que influyen en el contenido de la experiencia perceptual, mientras que el inferencialismo constituye una tesis acerca de cómo esa experiencia se genera o constituye. En un sentido próximo al fundherentismo de Haack (1993, 2014) experiencia y razones forman parte de una misma estructura justificativa, aunque cumplen funciones diferentes. Esto me permite sostener que la experiencia proporciona el punto de apoyo sobre el que se articulan las razones, sin que ello implique reducir unas a las otras. En otras palabras, a mi juicio, la experiencia perceptual constituye el punto de partida a partir del cual el razonamiento inferencial se despliega.

Esta distinción requiere algunas precisiones. El punto remite a una discusión más amplia entre posiciones inferencialistas y no inferencialistas, según si la experiencia perceptual es el producto de un proceso inferencial o no[18].

Ciertos enfoques inferencialistas y bayesianos de la percepción, sostienen que los estados perceptivos serían el producto de procesos inferenciales —incluso a nivel subpersonal— que integran información sensorial con creencias o experiencias previas (Siegel, 2017; McGrath, 2013). Intentaré ilustrar esta posición. Supongamos que usted abre la ventana y «ve» que está lloviendo. Desde un enfoque inferencialista esta experiencia podría entenderse como el resultado de un proceso, posiblemente inconsciente, mediante el cual el sistema perceptivo integra información sensorial (patrones visuales, movimiento de gotas, sonido, etc.) y genera un estado perceptivo con contenido determinado: está lloviendo. En este sentido, la experiencia perceptual seria análoga a una conclusión: el producto de una integración informada guiada por supuestos previos o probabilidades[19].

---

[18] Se trata de un debate amplio y profundamente desarrollado, cuya reconstrucción exhaustiva excede este espacio, por lo que aquí me limitaré a retomar aquellos aspectos que resultan relevantes para el problema que se analiza.

[19] El Premio Nobel Gerald Edelman, junto a Giulio Tononi, explican la experiencia consciente mediante la concurrencia de las propiedades de la integración y de la informatividad. Con la primera a

Retomemos el caso comentado al comienzo de la medica forense que observa manchas rojas (petequias) en los glóbulos oculares que un observador no entrenado sería incapaz de ver. Una reconstrucción inferencialista tendería a entender esa experiencia como el resultado de un proceso del razonamiento del siguiente tipo:

(i) Premisa 1: En los casos de estrangulamiento suelen observarse petequias en los glóbulos oculares.

(ii) Premisa 2: En el cadáver se presentan determinadas características visuales (pequeñas hemorragias puntiuniformes en los glóbulos oculares).

(iii) Conclusión perceptual: Hay petequias.

Y luego:

(iv) Premisa 3: La presencia de petequias es consistente con la hipótesis de estrangulamiento.

(v) Conclusión probatoria: La hipótesis de estrangulamiento recibe apoyo.

Bajo esta concepción, la primera conclusión «hay petequias» seria el resultado de una inferencia. Es decir, la experiencia perceptual misma tendría estructura de razonamiento.

Como advierte Teng (2020), este tipo de enfoques enfrenta dificultades tanto a nivel psicológico como epistemológico. Por un lado, no resulta claro que la experiencia perceptual tenga la estructura de una inferencia, esto es, que consista en una transición de estados mentales que funcione como un razonamiento articulado en términos de premisas y conclusión, susceptible de reconstrucción y evaluación en esos términos. Por el otro, su generalización puede conducir a un escepticismo excesivo respecto de la justificación perceptual, en la medida en que, si las experiencias fueran el resultado de inferencias, su fuerza justificatoria dependerá de la corrección de esos procesos. Dado que el sistema perceptivo funciona frecuentemente mediante heurísticas, supuestos implícitos o integraciones probabilísticas de la información, ello implicaría que una parte significativa de nuestras experiencias estarían epistémicamente degradadas, debilitando así su papel como fuente de justificación.

Las posiciones no inferencialistas, de la que forman parte Siegel, Stokes, Chudnoff, entre otros, niegan que ambas tesis sean equivalentes (que la experiencia perceptual al estar cognitivamente penetrada tenga estructura inferencial) y sostienen que la influencia de estados cognitivos sobre la experiencia no obliga a reconstruirla como una inferencia. En el enfoque que aquí se propone, esa misma experiencia no es interpretada como resultado de una inferencia, sino como el punto de partida del razonamiento. El entrenamiento y la experiencia modifican qué rasgos del entorno se vuelven perceptualmente salientes, de modo que la experiencia perceptual de la

---

lo que se apunta es a la unidad del estado perceptivo (no es posible dividir dicha experiencia en estados independientes). Con la segunda, en cambio, se pone énfasis en la diferenciación del estado experiencial en que nos encontramos con la infinidad de posibilidades que le compiten (Edelman y Tononi, pp. 32-48).

médica forense consiste en «ver petequias». Las inferencias aparecen posteriormente, cuando esa experiencia es articulada con generalizaciones médicas y conectadas con la hipótesis de estrangulamiento. En este sentido, la percepción experta constituye el punto de partida sobre el que se construyen las cadenas inferenciales que permiten justificar y comunicar las conclusiones alcanzadas.

Esta distinción no es solo descriptiva, sino normativa. Mientras que en el primer caso la fuerza justificatoria de la experiencia dependería de la corroboración del proceso inferencial que la produce, en el segundo la experiencia funciona como base epistémica desde la cual se articulan las inferencias, sin ser ella misma evaluada como una inferencia en sentido estricto.

Dado que los expertos poseen pericial perceptual, ese punto de partida no es neutro, pues se vuelve más rico y estructurado, aunque también más opaco a terceros que no poseen el mismo nivel de dominio técnico y experiencia. Esto plantea un problema, ¿dónde se ubica el escrutinio epistémico en contextos como el jurídico?

En términos generales, cuando los tribunales evalúan la percepción de un testigo, lo que hacen es controlar las condiciones externas bajo las cuales la observación tuvo lugar, tales como la distancia, la iluminación o la posibilidad material de advertir ciertos rasgos del entorno. En este sentido, las afirmaciones acerca de lo percibido pueden ser puestas en entredicho cuando existen razones para pensar que el sujeto no estaba en condiciones de percibir aquello que declara. El escrutinio de la percepción recae, así, principalmente sobre las circunstancias en que la observación se produjo.

En el caso de la experticia, esta evaluación se complejiza, en la medida en que el punto de partida perceptual incorpora estructuras cognitivas especializadas. En este nivel, existe un aspecto de la experiencia perceptual que no parece susceptible de poder controlarse, pues el tribunal no puede evaluar qué rasgos se vuelven perceptualmente salientes para el experto ni cómo estos organizan la experiencia a la luz de categorías y conocimientos especiales. Así, por ejemplo, no resulta razonable exigir que la jueza «vea» las petequias, siquiera del mismo modo que la médica forense. Lo que sí se puede examinar es si, dadas las condiciones del caso, era posibles observarlas. Entonces, el tribunal puede evaluar si la experta se encontraba en condiciones adecuadas para percibir aquello que declara, pero no puede exigir una reproducción de la experiencia perceptual misma. Ello se debe, por un lado, a que ciertos rasgos del entorno adquieren una saliencia perceptual que no está igualmente disponible para el novato, y por el otro, a que la afirmación «hay petequias» incorpora categorías y lenguaje técnico que permite describir y clasificar aquello que es percibido. En otras palabras, se puede controlar las condiciones de observación, pero no lo que se vuelve saliente y la manera en que se organiza la percepción.

Si se asume que la percepción funciona como punto de partida del razonamiento, entonces las dificultades asociadas a la percepción experta se proyectan, al menos parcialmente, sobre las inferencias que se construyen a partir de ella. En este sentido, la opacidad que caracteriza a la experiencia perceptual no desaparece en el nivel inferencial, puesto que las cadenas de razonamiento se apoyan en un punto de partida

que no es plenamente accesible para los novatos ni los tribunales. Esto podría llevar a pensar entonces que, en el caso de los expertos, tanto su percepción como las inferencias que se formulan a partir de ella, escapan al control epistémico. Sin embargo, un argumento de este tipo parece no conciliar nuestras prácticas ordinarias y jurídicas de evaluación del conocimiento, en el que el escrutinio se dirige a examinar las conexiones que se establece entre lo que se ve y lo que se dice.

Una estrategia para evitar cierto escepticismo respecto de la justificación perceptual es advertir que, en el nivel inferencial, el tribunal posee una ventaja que no tiene en la estructura perceptual: las inferencias deben ser comunicadas, explicitadas, contrastadas y sometidas a control intersubjetivo. Esto sugiere que el control epistémico opera por grados, en la medida en que las actividades involucradas (percepción y razón) representan distintos niveles de explicitación. Es decir, aun cuando la opacidad asociada a expertise perceptual se proyecte en el nivel inferencial, las inferencias ofrecen mayores posibilidades de explicitación y, por ende, de evaluación crítica.

Concebir la experiencia perceptual como el un punto de partida epistémico que proporciona razones *prima facie* para la formación de creencias, no implica adoptar un dogmatismo ingenuo, según el cual «la percepción siempre justifica»; sino reconocer que la percepción posee valor epistémico aun cuando está cognitivamente penetrada, es susceptible a sesgos y no resulta completamente transparente ni para el propio sujeto ni para terceros. Por esta razón, su justificación no descansa exclusivamente ni en la experiencia perceptual ni en las inferencias que se construyen a partir de allí, sino en una relación entre ambas.

Esta forma de comprender la relación entre percepción y razones es precisada a partir de Haack (2020), quien describe la prueba de afirmaciones empíricas como un entramado en el que la experiencia y las razones —incluidas las creencias de fondo— operan conjuntamente, «como pistas y entradas entrecruzadas que se ramifican como en un crucigrama» (p. 136). Lejos de concebir a la percepción como una inferencia o de aislarla como una fuente autosuficiente de justificación, esta imagen permite dar cuenta de su papel como elemento que se integra en una red más amplia de apoyo epistémico.

En el caso del conocimiento experto, dicho entramado se inicia en una experiencia perceptual que no es inferencialmente derivada, pero cuya fuerza justificativa depende de cómo se articula con las cadenas inferenciales que la hacen comunicable y susceptible de evaluación. Si bien estas últimas heredan parte de la obscuridad asociada a la percepción especializada, poseen mayores posibilidades de explicitación y, por ello, ofrecen un espacio privilegiado —aunque no exclusivo— para el escrutinio epistémico. Así, el control puede operar gradualmente en función del grado de explicitación de los distintos componentes del razonamiento experto, sin por ello quedar exento de limitaciones.

En este sentido, la diferencia entre novatos y expertos involucra una diferencia estructural en la forma de percibir y correlativamente el tipo de justificación que

están en condiciones de ofrecer. Las creencias del experto se originan en experiencias perceptuales que capturan aspectos del mundo no accesibles al observador común, lo que ubica su testimonio en una posición epistémica diferenciada y explica tanto su mayor fuerza justificadora como los riesgos específicos que la acompañan.

A modo de cierre, retomo el caso de la médica forense para precisar este aspecto central. Cuando la experta declare ante el tribunal, la brecha perceptual entre expertos y novatos reaparecerá en forma epistémica: no solo difiere en lo que saben, sino en lo que son capaces de percibir, y esa diferencia incide directamente en el tipo de razones que están en condiciones de ofrecer. En una primera instancia, la perita informará una observación perceptual: la presencia de petequias en los glóbulos oculares. En una segunda instancia, establecerá la conexión inferencial que permite dotar de relevancia ese dato, explicando por qué ese hallazgo es consistente con la hipótesis de estrangulamiento.

Esta doble dimensión del testimonio experto[20] sugiere algunas pistas respecto de los problemas del control epistémico. Quizás una parte de las dificultades de la valoración del conocimiento experto ha tendido a pasar inadvertida porque la atención se ha concentrado, principalmente, en la formulación de inferencias y no tanto en el origen del conocimiento ni en las formas en que este se estructura en la experiencia perceptiva especializada que le sirve de base. Si es ello es así, cabría pensar que las propias inferencias expertas —en la medida en que se apoyan en una percepción parcialmente inaccesible para terceros— también se encuentra sujeta a límites en su control más profundas o extendidas de lo que habitualmente se supone. Precisar el alcance y las consecuencias de esta diferencia sea, quizás, una cuestión que merece una reflexión más detenida. He intentado aquí ofrecer algunas claves para pensar este problema bajo una comprensión más amplia de la valoración de la prueba pericial, poniendo de relieve ciertos límites y dificultades que suelen permanecer implícitos y abrir una línea de reflexión acerca de sus implicancias para el derecho probatorio.

6. CONCLUSIONES

Este artículo partió del interrogante acerca de si la diferencia entre novatos y expertos se reduce a una cuestión de grados de conocimiento o si involucra diferencia en la forma en que se configura la experiencia perceptiva. A lo largo del trabajo se ha sostenido que la experticia no solo implica disponer de más conocimiento, sino

[20] Esta doble dimensión del testimonio experto ha sido destacada en la literatura. Como señalan Wahlberg y Dahlman (2021), el perito no se limita a actuar como un testigo que informa sus observaciones, sino que también opera en un nivel experto, en el que evalúa y articula esas observaciones a la luz de conocimientos especializados. Ambas funciones —la testimonial y la pericial— no solo coexisten, sino que suelen encontrarse estrechamente entrelazadas en la práctica pericial. Esta caracterización resulta consistente con la distinción aquí propuesta entre un nivel perceptual, que funciona como punto de partida, y un nivel inferencial, en el que ese contenido adquiere relevancia probatoria.

de percibir de manera distinta, en virtud de la incidencia de estados cognitivos que organizan el contenido de la experiencia.

En este marco, y a partir de la noción de penetrabilidad cognitiva, se mostró que la percepción experta constituye una forma de experiencia en la que lo sensorial y cognitivo se encuentran estrechamente entrelazados. Esta caracterización permite dar cuenta tanto de su valor epistémico como de las particularidades que presenta en contextos de evaluación intersubjetiva.

Sobre esta base, se introdujo una distinción entre la percepción como punto de partida del razonamiento y las cadenas inferenciales mediante las cuales los expertos articulan y comunican sus conclusiones. A partir de ello, se argumentó que la mayor robustez del conocimiento experto no se explica únicamente por la disponibilidad de conocimiento proposicional, sino también por la calidad de los enlaces inferenciales que se construyen a partir de la experiencia.

Finalmente, se examinó cómo esta distinción incide en el control epistémico en contextos jurídicos. Se sostuvo que la evaluación del conocimiento experto no se dirige de manera uniforme a todos los componentes del testimonio, sino que adopta formas diferenciadas y graduales, en función del grado de explicitación de cada uno de ellos. Mientras que la experiencia perceptual especializada solo puede ser considerada a partir de las condiciones en que se produce la observación, las cadenas inferenciales ofrecen mayores posibilidades explicitación y control intersubjetivo. Sin embargo, las dificultades asociadas a la percepción experta persisten en el nivel inferencial. En este sentido, el trabajo sugiere que una parte de los problemas asociados a la valoración del conocimiento experto podría encontrarse no solo en las inferencias, sino también en las formas en que el conocimiento se estructura en la experiencia perceptual que le sirve de punto de partida.

BIBLIOGRAFÍA

Accatino, Daniela. (2018). La prueba pericial y la epistemología del testimonio: ¿Una teoría y dos modelos? *Anuario de Filosofía Jurídica y Social, 34*, 5–17.

Burge, Tyler. (2010). *Origins of objectivity*. Oxford University Press.

Chudnoff, Elijah. (2019). Experience and epistemic structure: Can cognitive penetration result in epistemic downgrade? En Jonathan Genone, Tania Lombrozo y Irving Konečni (Eds.), *Inference and consciousness* (pp. 255–274). Routledge.

Dror, Itiel E. (2020). Cognitive and human factors in expert decision making: Six fallacies and the eight sources of bias. *Analytical Chemistry, 92*(12), 7998–8004. *https://doi.org/10.1021/acs.analchem.0c00704*

Dror, Itiel E., McCormack, Bridget, & Epstein, Jules. (2015). Cognitive bias and its impact on expert witnesses and the court. *Judges' Journal, 54*, 8–12.

Edelman, Gerald, & Tononi, Giulio. (2005). *El universo de la conciencia: Cómo la materia se convierte en imaginación*. Crítica.

Gascón Abellán, Marina. (2013). Prueba científica: Un mapa de retos. En Carmen Vázquez (Ed.), *Estándares de prueba y prueba científica. Ensayos de epistemología jurídica* (pp. 181–202). Marcial Pons.

Gascón Abellán, Marina. (2019). Prevención y educación: El camino hacia una mejor ciencia forense en el sistema de justicia. En Jordi Ferrer Beltrán & Carmen Vázquez (Eds.), *El razonamiento probatorio en el proceso judicial. Un encuentro entre diferentes tradiciones* (pp. 237–265). Marcial Pons.

Gelfert, Axel. (2014). *A critical introduction to testimony*. Bloomsbury.

González Lagier, Daniel. (2018). Tres modos de razonar sobre hechos (y algunos problemas sobre la prueba judicial planteados a partir de ellos). En Carmen Vázquez (Coord.), *Hechos y razonamiento probatorio* (pp. 17–43). Editores del Sur.

González Lagier, Daniel. (2022). Inferencia probatoria y valoración conjunta de la prueba. En Jordi Ferrer Beltrán (Coord.), *Manual de razonamiento probatorio* (pp. 353–396). Suprema Corte de Justicia de la Nación. *https://www.scjn.gob.mx/derechoshumanos/sites/default/files/Publicaciones/archivos/2024-02/Manual%20de%20razonamiento%20probatorio_0.pdf*

Hanson, Norwood Russell. (1977). *Observación y explicación: Guía de la filosofía de la ciencia*. Alianza.

Haack, Susan. (1993). *Evidencia e investigación: Hacia la reconstrucción en epistemología*. Tecnos.

Haack, Susan. (2014). *Evidence matters: Science, proof, and truth in the law*. Cambridge University Press.

Haack, Susan. (2020). *Filosofía del derecho y de la prueba: Perspectivas pragmatistas*. Marcial Pons.

Herdy, Rachel. (2020). Ni educación ni deferencia ciega: Hacia un modelo crítico para la valoración de la prueba pericial. *Discusiones, 24*(1), 87–112.

Jenkin, Zoe, & Siegel, Susanna. (2015). Cognitive penetrability: Modularity, epistemology, and ethics. *Review of Philosophy and Psychology, 6*(4), 531–545.

McDowell, John. (1994). *Mind and world*. Harvard University Press.

McDowell, John. (2009). Avoiding the myth of the given. En *Having the world in view* (pp. 256–272). Harvard University Press.

McGrath, Matthew. (2013). Phenomenal conservatism and cognitive penetration: The "bad basis" counterexample. En Chris Tucker (Ed.), *Seemings and justification: New essays on dogmatism and phenomenal conservatism* (pp. 225–247). Oxford University Press.

Páez, Andrés. (2026). La negligencia institucional frente al riesgo epistémico de los sesgos cognitivos en el derecho. *Questio Facti, 10*.

Pereira, Francisco. (2021). *Ver no es creer: Sobre el rol de los conceptos en la experiencia visual*. Gedisa.

Ransom, Madeleine. (2020). Expert knowledge by perception. *Philosophy, 95*(3), 309–335.

Sellars, Wilfrid. (1956). Empiricism and the philosophy of mind. En Herbert Feigl & Michael Scriven (Eds.), *Minnesota studies in the philosophy of science* (Vol. 1, pp. 253–329). University of Minnesota Press.

Siegel, Susanna. (2009). The epistemic impact of the etiology of experience. *Philosophical Studies, 162*(3), 697–722.

Siegel, Susanna. (2010). *The contents of visual experience*. Oxford University Press.

Siegel, Susanna. (2012). Cognitive penetrability and perceptual justification. *Noûs, 46*(2), 201–222.

Siegel, Susanna. (2015). Epistemic charge. *Proceedings of the Aristotelian Society, 115*(3), 277–306.

Sosa, Ernest. (2007). *A virtue epistemology: Apt belief and reflective knowledge* (Vol. 1). Oxford University Press.

Stokes, Dustin. (2021). On perceptual expertise. *Mind & Language, 36*(2), 241–263.

Teng, Lu. (2020). Cognitive penetration: Inference or fabrication? *Australasian Journal of Philosophy*. Advance online publication. *https://doi.org/10.1080/00048402.2020.1812095*

Torretti, Roberto. (2012). *Inventar para entender. Creatividad e innovación: Raíces de la explicación científica*. Ediciones Universidad Diego Portales.

Toribio, Josefa. (2007). Contenido no conceptual. *Philosophy Compass, 2*(3), 445–460.

Vázquez, Carmen. (2015). *De la prueba científica a la prueba pericial*. Marcial Pons.

Vázquez, Carmen. (2020). La división del trabajo cognitivo en la prueba pericial: El juez, las partes, los peritos y las comunidades expertas. *Discusiones, 24*(1), 113–140.

Wahlberg, Lena, & Dahlman, Christian. (2021). The role of the expert witness. En Christian Dahlman, Alex Stein & Gerd Gigerenzer (Eds.), *Philosophical foundations of evidence law* (pp. 53–66). Oxford University Press.

Williams, Bernard. (2002). *Truth and truthfulness*. Princeton University Press.

CONJETURAS
Y REFUTACIONES

Quaestio facti. Revista Internacional sobre Razonamiento Probatorio
Quaestio facti. International Journal on Evidential Legal Reasoning
Sección: Conjeturas y refutaciones
2026 I 11 pp. 143-156
Madrid, 2026
DOI: 10.33115/udg_bib/qf.i11.23268
Marcial Pons Ediciones Jurídicas y Sociales
© Silvia Bozza
© Franco Taroni
© Colin Aitken
ISSN: 2604-6202
Recibido: 06/02/2026 | Aceptado: 07/05/2026 | Publicado *online*: 29/05/2025
Editado bajo licencia Reconocimiento 4.0 Internacional de Creative Commons

IN DEFENCE OF SUBJECTIVITY: EXTENDING THE ARGUMENT FOR A TRANSPARENT TRIAL. (A SUPPORT ON "BAYESIAN MODELLING OF CRIMINAL CASES AS A WHOLE. A PHILOSOPHICAL REFLECTION ON DUTCH CASE LAW")

Silvia Bozza*
Franco Taroni**
Colin Aitken***

ABSTRACT: This comment addresses the "prior challenge" in forensic Bayesian modelling, recently highlighted by Anne Ruth Mackor (2026). We argue that the perceived lack of frequency data is not an insurmountable obstacle but a misconception rooted in an outdated view of probability. By adopting a radical subjectivist perspective based on de Finetti's teachings, we reframe probability as a coherent representation of a decision-maker's state of knowledge. We advocate for a strict functional separation: forensic experts provide the likelihood ratio based on technical findings, while the court assigns prior odds based on the specific case context. Through sensitivity analysis, we demonstrate that the subjectivity of priors can be viewed not as a source of arbitrariness but rather as a transparent and auditable mechanism that enhances judicial accountability. Ultimately, the Bayesian model is presented as a logical necessity for preventing miscarriages of justice.

KEYWORDS: bayesian modelling; forensic science; subjective probability.

* Department of Economics, Ca' Foscari University of Venice, Venice, Italy. School of Criminal Justice, University of Lausanne, Lausanne, Switzerland. E-mail: *Silviab@unive.it*, *Silvia.Bozza@unil.ch*.

** School of Criminal Justice, University of Lausanne, Lausanne, Switzerland. E-mail: *Franco.Taroni@unil.ch*.

*** School of Mathematics and the Maxwell Institute of Mathematical Sciences, The University of Edinburgh, Edinburgh, United Kingdom. E-mail: *C.Aitken@ed.ac.uk*.

SUMMARY: 1. INTRODUCTION: REFRAMING THE CHALLENGE OF THE CHOICE OF PRIORS.— 2. WHAT IS SUBJECTIVE PROBABILITY? A BRIEF CLARIFICATION: 2.1. Probability as a state of knowledge; 2.2. The role of frequencies; 2.3. How to quantify subjective beliefs: the logic of comparison.— 3. THE ONTOLOGICAL STATUS OF PROBABILITY AND THE MYTH OF "OBJECTIVE" DATA.— 4. DEFINING THE BOUNDARIES (TO EACH ONE THEIR OWN PROBLEM): 4.1. The role of the expert: the likelihood ratio; 4.2. The role of the judge: the prior and the context; 4.3. Avoiding the contextual vacuum.— 5. ROBUSTNESS THROUGH SENSITIVITY ANALYSIS (FROM THEORY TO JUDICIAL PRACTICE): 5.1. The "Washing out" of the Prior; 5.2. Transparency as a shield against arbitrariness.— 6. THE DANGER OF THE CONTEXTUAL VACUUM AND THE PREVENTION OF MISCARRIAGES OF JUSTICE: 6.1. Context as a logical necessity; 6.2. Misuse vs. invalidity.- 6.3. Bridging the gap: from data to justice.— 7. CONCLUSION: EMBRACING COHERENCE IN LEGAL FACT-FINDING.— ACKNOWLEDGEMENTS.— REFERENCES.

1. INTRODUCTION: REFRAMING THE CHALLENGE OF THE CHOICE OF PRIORS

The application of Bayesian networks and probabilistic modelling to entire criminal cases represents one of the most significant shifts in contemporary legal epistemology. In her comprehensive reflection on Dutch case law, Anne Ruth Mackor (2026) provides a crucial map of this territory, exploring whether the Bayesian model can truly provide guidance for the judicial evaluation of evidence and interpretation of cases; an aspect already extensively discussed more than forty years ago (see, e.g. Eggleston, 1983; Lempert, 1977; Tillers, 1986; Tillers and Green, 1988; Kaye, 1988). While acknowledging its power as a rational method (Mackor, 2026 at p. 1), Mackor identifies a recurring friction point that often halts the model's adoption: the assignment of prior probabilities.

As Mackor correctly points out, the transition from forensic laboratories, where the likelihood ratio (LR for short) measure is the standard output, to the courtroom is hindered by the lack of relevant data. She notes that in criminal law, there are usually no reliable statistics about the frequencies of the events that are relevant to determine the prior odds and, consequently, such assignments must often be a matter of subjective estimation (Mackor, 2026 at p. 4)[1]. For many legal and scientific scholars and practitioners, this subjectivity is viewed with suspicion, as if it introduced an arbitrary element into an otherwise rigorous scientific process.

The purpose of this comment is to challenge the notion that subjective assignment is a weakness of the Bayesian system. We argue that the so-called "prior challenge" is essentially a non-problem if one adopts a radical subjectivist perspective

[1] Note that the term "estimation" can be misleading. As discussed by Fischhoff and Beyth-Marom (1983), 'The term "assign" is used rather than "estimate" to emphasize that a probability expresses one's own feelings rather than an appraisal of a property of the physical world. Thus, there is no "right" probability value for a particular statement' (at p. 240).

rooted in the teachings of Bruno de Finetti (de Finetti, 1968, 1972) and Dennis Lindley (Lindley, 1991). Subjectivity, in this sense, is not a synonym for arbitrariness, but rather a transparent quantification of a decision-maker's state of knowledge.

It is important to note that while Mackor (2026) rightly emphasizes the role of Bayesian networks in legal fact-finding, the present comment does not delve into their structural implementation. We fully acknowledge that Bayesian networks represent the most robust and coherent framework for evaluating complex patterns of evidence and their inferential interactions, a field where we have contributed extensively in our own previous work (e.g., Aitken and Gammerman, 1989; Taroni *et al.*, 2014) following the teaching of David Schum and Jay Kadane (Schum, 1994; Kadane and Schum, 1996). However, the focus of this paper is not on the architecture of evidence evaluation, but on the epistemological hurdle that precedes it: the prior challenge. By addressing the nature of subjective probability and the division of labour between the expert and the judge, we aim to clear the ground, ensuring that the powerful tools offered by Bayesian networks are built upon a firm and transparent logical foundation.

The present article is organized as follows. After this introduction, Section 2 provides a foundational definition of subjective probability, clarifying that it represents a state of knowledge rather than a physical property of the world, and illustrating how it can be rigorously quantified. Section 3 addresses the ontological nature of probability, arguing that the lack of frequency-based statistics does not preclude the formulation of coherent prior odds. Section 4 explores the division of labour between the forensic expert and the judge, responding to Mackor's queries regarding *who* should use the model (Mackor, 2026 at p. 11) by emphasizing that the prior odds is the exclusive and necessary domain of the court. Section 5 addresses common critiques regarding the subjectivity of priors by employing sensitivity analysis; this demonstrates that the Bayesian framework remains robust even in the absence of precisely defined prior distributions. Finally, Section 6 warns against the dangers of attempting to evaluate evidence in what may be thought of as a contextual vacuum, concluding that the Bayesian model is not just an option, but a logical necessity for preventing miscarriages of justice.

2 WHAT IS SUBJECTIVE PROBABILITY? A BRIEF CLARIFICATION

To understand why the prior challenge is a manageable hurdle, one must first clarify the nature of probability itself. As it has been argued extensively in philosophical and statistical literature (e.g. de Finetti, 1974), there is a persistent misconception in legal scholarship that for a probability to be so-called "reliable", it must be "objective" (in some sense), meaning it must be derived from long-run frequencies or hard statistical data.

2.1. Probability as a state of knowledge

Following the radical subjectivist tradition of Bruno de Finetti, we maintain that probability is not an external, physical property of the world (like mass or length), but a representation of an individual's degree of belief in the occurrence of an event, given the information available at that time. In this sense, all probabilities are conditional: they depend on the state of knowledge (at a given time) of the person making the assessment.

For a judge, this is not a revolutionary concept; it is the formalization of what legal systems call "free evaluation of evidence". When a judge assesses a case, they are updating their personal state of uncertainty based on the evidence presented.

2.2. The role of frequencies

A common critique is that subjective probability ignores what is called "hard data". On the contrary, the subjectivist framework perfectly reconciles personal beliefs with frequencies (Lindley, 1991). As shown in numerous works on de Finetti's Representation Theorem and exchangeability (see, e.g., Dawid and Galavotti, 2009; Taroni *et al.*, 2018), frequencies are simply a specific type of information that a rational agent uses to update their beliefs.

Data does not speak for itself; it requires a subject to interpret its relevance to the specific, unique case at hand. Therefore, subjectivity is not a source of error, but the necessary mechanism through which evidence is given meaning (Press and Tanur, 2001). By naming it *subjective*, we are not admitting to a lack of rigor, we are providing a transparent account of the reasoning process, making it auditable and logically coherent.

2.3. How to quantify subjective beliefs: the logic of comparison

A frequent objection is: 'How can a judge assign a precise number to a subjective belief?' The answer lies not in a guess, but in a process of logical comparison (Biedermann *et al.*, 2013, 2016). As many quarters have discussed in scientific and legal literature (see, e.g., Lindley, 1975), quantifying a subjective probability means expressing a preference between different uncertain outcomes.

A common method is the reference experiment (or the betting analogy): if a judge says the probability of a hypothesis is 10%, they are saying they find that hypothesis as likely as drawing a red ball from an urn containing 10 red balls and 90 white ones. This mental exercise forces the decision-maker to calibrate their internal uncertainty against a known external standard.

For sake of illustration, consider the "prior" in a burglary where a DNA trace is found. Before looking at the DNA result, the judge must assess the "prior" probability that the suspect is the source.

— If the suspect was found 500 metres from the scene shortly after the crime with no alibi, the judge might reasonably assign a numerical value to their degree of belief (e.g., 1/100) based on the pool of potential people in that area at that time.

— If the suspect was instead identified only via a national database search with no other link to the city, the prior would be much lower (e.g., one over the size of some relevant population, where relevance has to be determined with reference to other information surrounding the case).

The number is not arbitrary because it must be coherent with the available background information. If the judge's assessment is transparently stated, it can be debated and challenged by the parties. Subjectivity thus becomes an instrument of precision: it allows the court to distinguish between a vague suspicion and a grounded hypothesis using the universal language of logic.

3 THE ONTOLOGICAL STATUS OF PROBABILITY AND THE MYTH OF "OBJECTIVE" DATA

In her analysis, Mackor (2026) identifies a fundamental hurdle for the legal application of Bayesian networks: the perceived lack of reliable statistics to determine prior odds. She notes that, in the absence of such data, the assignment of probabilities becomes a matter of subjective estimation, which may lead to a lack of intersubjective agreement among the actors in the trial. This challenge highlights a widespread epistemological friction in the forensic field: a lingering frequentist-like expectation that probability should ideally be an objective property of the world, measurable only through repeated trials. From this common viewpoint, which often conflates "subjective" with "arbitrary", subjectivity is seen as a fallback to be used only when so-called hard data are missing.

However, as we have argued in a previous publication (Taroni *et al.*, 2018), this objection stems from a misunderstanding of the nature of probability itself. Following the subjectivism of Bruno de Finetti, we maintain that "Probability does not exist" (de Finetti, 1974) as an external, physical entity[2]. Instead, probability is a formal expression of a subject's degree of belief based on the information available. Far from being a sign of arbitrariness, the subjectivist approach simply provides a formal, transparent structure to the reasoning processes that judges already exercise daily. By

[2] As Bruno de Finetti (1974) provocatively stated in his seminal essay, the concept of probability is not an objective property of the physical world, but a formal expression of a subject's state of knowledge. For the legal evaluator, this shift is crucial: it moves the focus from the search for an elusive "objective frequency" to the construction of a coherent and transparent belief system based on available information.

explicitly quantifying uncertainty, the judge shifts from an opaque, unchallengeable "gut feeling" to clear lines of reasoning.

To clarify this point, let us consider two examples that illustrate why what is believed to be "objective" statistics are, in fact, always filtered through subjectivity:

1) The DNA database example: consider the occurrence of a genetic profile (e.g., 1 in a million) derived from a database. This number is a relative frequency, a historical record of past observations. To use this as a probability for the case at hand, the evaluator must assume that the database population is one of which the defendant could be considered a member. If the defendant (and the real donor of the recovered stain) belongs to a specific ethnic subgroup not represented in the database, the statistic becomes irrelevant. The choice to use that 1-in-a-million figure is, therefore, a subjective judgment based on the assumption of symmetry.

2) The local crime rate example: if a judge knows that 80% of burglaries in a specific district are committed by local residents, this 0.8 frequency does not automatically become the prior probability for a new suspect. The judge must decide if the statistic (80%) about burglaries in a specific district are committed by local residents is relevant for the specific circumstances of the current arrest (e.g., the suspect was caught at a train station). Without this subjective leap, the figure 80% is irrelevant.

Analogously, see also Lindley (2014 at p. 150): 'The records of the doctor observing the presence or absence of a symptom with a disease, you might think exchangeable,[3] though if you knew the sexes of the patients and thought the disease was sex-related, you might not. This example also serves to illustrate an important point, that since the definition of exchangeable depends on your probabilities, it depends on your knowledge base, and a series exchangeable under one base, without knowledge of sex, may fail to be under another, with knowledge of sex.'

As we would like to express, the subjectivist framework does not ignore empirical data; it provides the only coherent way to integrate them through the Representation Theorem (de Finetti, 1982; Dawid, 2004). Frequency data are simply observations of past events. For these observations to inform a current, unique case, a judge must assume a degree of symmetry (or exchangeability) between those past events and the case at hand.

As Kadane (1995) has argued, the requirement for a prior probability is not that it be objectively true, but that it be coherent. By requiring the judge to assign a prior probability, the Bayesian framework forces an unprecedented level of transparency. It compels the court to state its initial assumptions clearly, rather than leaving them buried in the black box of unquantified judicial intuition. In this sense, the subjec-

[3] "Exchangeable" in the sense that the order in which the records are considered is not relevant to any inference which may be made from them. Exchange the order of consideration and the inference remains the same. If there is knowledge of sex this may not be the case.

tive "estimation" Mackor (2026) refers to is actually a safeguard: it transforms hidden biases into explicit and testable situations. The subjectivist approach offers a coherent language and tool set for handling uncertainty.

4 DEFINING THE BOUNDARIES (TO EACH ONE THEIR OWN PROBLEM)

A central question raised by Mackor (2026) concerns the practical implementation of the Bayesian perspective: should non-forensic experts and judges engage directly with probabilistic reasoning? She notes that while the Netherlands Forensic Institute (NFI) provides likelihood ratios as a coherent metric to assess the value of scientific findings, the integration of these results into a broader criminal case remains a challenge for the judiciary (at p. 12). From our perspective, this difficulty does not stem from the complexity of the Bayesian approach itself, but from a persistent lack of clarity regarding the division of labour between the expert and the trier of fact. The most effective way to address the so-called "prior challenge" is to adhere to a strict functional separation: to each one their own problem to solve!

Guidance in the use of Bayes nets for judges, lawyers, forensic scientists and expert witnesses is given in the third of four guides on *Communicating and Interpreting Statistical Evidence in the Administration of Criminal Justice* published by the Royal Statistical Society (Roberts and Aitken, 2014). There it is noted that "Bayesian networks assist their users … to understand the structure of complex inferential problems, to form a better appreciation of mutual dependencies between uncertain events and compound probabilities, and to express this understanding in a graphical form that both assists in deepening their own comprehension and enables them to communicate their insights to others. Bayes nets help to clarify the nature of arguments predicated on probabilistic argument and thus promote logical analysis and rational further discussion and evaluation of factual propositions".

Bayes nets provide an answer to both parts of Mackor's second question. They can be used by (a) an expert and (b) the court itself for "integral modelling of a case as a whole".

Bayes nets can be used in the investigative stage of a crime (Aitken *et al.*, 1996). Factors arising in the investigation can be incorporated into a graphical structure and their inter-relationships clarified visually. Similarly, a graphical structure can be used by advocates for assistance in summarising the case for the judge (e.g. Edwards, 1991). Then assumptions and probabilistic assignations can easily be checked by the judge for their reliability in assistance for the final decision about the outcome of the case.

4.1. The role of the expert: the likelihood ratio

The forensic scientist's responsibility is to evaluate the scientific findings (or the strength of the association observed between control and questioned materials) by comparing the probability of the findings under (at least) two competing hypotheses. This results in a metric called the likelihood ratio (ENFSI, 2015), which we consider a logical necessity rather than a mere professional convention. In this stage, the scientist must remain within the bounds of their technical expertise. This necessity arises from the requirement to maintain a coherent inferential process while respecting the functional separation between the expert's evaluation of evidence and the court's assessment of the case. They do not, and should not, provide the prior odds, as they do not have access to the full context of the case described by what is known as background information. If an expert were to incorporate prior probabilities into their report, they would be encroaching upon the judicial prerogative, effectively deciding the case rather than evaluating the findings (Biedermann *et al.*, 2007; Taroni and Biedermann, 2005).

The stricture that the scientist should not provide the prior odds does not mean the scientist cannot use a Bayesian network, with a node for the hypothesis of interest, to assist their evaluation of the evidence. The likelihood ratio is the ratio of the posterior odds in favour of a hypothesis to the prior odds. Thus, propagation of the evidence in a Bayes net will change the prior odds to posterior odds. The ratio of these two odds is the likelihood ratio. The value chosen for the prior odds cancels out with the posterior odds; increase the prior odds by a factor of 10, for example and the posterior odds will increase by a factor of 10 also, the likelihood ratio will be unchanged.

The professional guidelines for such reporting are clearly defined by the European Network of Forensic Science Institutes (ENFSI, 2015) and other scientific references (Berger, 2025), which mandate that forensic evaluations be presented as a measure of evidential strength rather than as a statement of guilt. Within this framework, the likelihood ratio is the most coherent measure for hypothesis confirmation, as demonstrated by philosophical and scientific literature (e.g. Taroni *et al.*, 2021). In addition, the likelihood ratio satisfies the necessary logical criteria for normative adequacy, offering a robust numerical expression for the impact of scientific findings on competing judicial hypotheses (Robertson *et al.*, 2016; Aitken *et al.*, 2021).

Unlike alternative, often informal, expressions of the value of evidence, the likelihood ratio provides a standardised, auditable metric that resists the logical pitfalls frequently encountered in courtrooms. By adhering to these rigorous reporting guidelines, the forensic scientist provides the court not with an opinion on the final outcome, but with a precise, logically sound tool to update their belief in light of the new scientific information (Thompson *et al.*, 2018).

It is essential to distinguish between the inferential stage, where we reason about the value of scientific findings and the decision-making stage, where the court must

decide on the ultimate issue, such as a defendant's guilt or innocence. Conflating these two stages is a primary source of logical incoherence. The Bayesian framework does not aim to automate the judge's decision, but to provide the necessary logical scaffolding to ensure that the transition from inference to decision is transparent. By formalizing one's degree of belief, we do not ignore uncertainty; rather, we provide a coherent language to manage it, ensuring that final decisions are not based on hidden intuition but on a structured, auditable reasoning process.

Entrusting the forensic scientist with the assignment of prior odds would effectively mean delegating the adjudication of the case to a laboratory. The expert's task is to provide the value of the evidence; the judge's task is to integrate this evidence into the broader context of the case. This separation of powers is the essential guarantee of a judicial system.

4.2. The role of the judge: the prior and the context

The judge (and in the adversarial system, a jury), on the other hand, is the only actor in the trial who possesses the integral view of the case. Therefore, the assignment of prior odds is the exclusive and necessary domain of the court. Mackor (2026) highlights difficulties in the determination as to how pieces of evidence should be combined (at p. 14-15, footnote 36). We argue that this difficulty is exacerbated when judges try to be seen to be objective by ignoring their own initial degrees of belief.

By using the odds form of Bayes' Theorem, the judge's role becomes clear: they take the likelihood ratio provided by the forensic expert and multiply it by their own prior odds, based on the context of the case, to arrive at the posterior odds. The context, or background information, of the case cannot be ignored. The Bayesian framework does not force a judge to become a statistician; rather, it provides a coherent language to ensure that the judge's integration of evidence is logically consistent.

4.3. Avoiding the contextual vacuum

The fear that judges are not ready for Bayesian modelling often leads to the dangerous practice of evaluating evidence in a vacuum. However, a piece of evidence has no meaning without a context; this is an established point dating back to Bertillon's 1886 and Locard's 1940 publications echoed by Robertson and Vignaux (1993). For instance, a DNA correspondence between genetic material associated with a crime scene and genetic material from a person of interest (offering a high value for the likelihood ratio if evidence is assessed under what are known as "source-level hypotheses") has different probative implications if the suspect was found near the crime scene (high prior) as opposed to cases where the suspect was identified through a random database search (low prior).

In conclusion, the Bayesian model is not an additional burden for the judge, but a tool for transparency. Its use clarifies the separation of the provision of the value of the evidence by the forensic scientist and the provision of the context by the judge. By separating these two components, the integrity of the trial is safeguarded from pitfalls of intuition (e.g. the prosecutor's fallacy) and it is ensured that the final inference and subsequent decision are a product of both scientific rigour and judicial responsibility.

5. ROBUSTNESS THROUGH SENSITIVITY ANALYSIS (FROM THEORY TO JUDICIAL PRACTICE)

A common anxiety among legal scholars is that different judges might assign different prior odds to the same case, leading to inconsistent or arbitrary outcomes. If the choice of prior is met with subjective assignment, how can the law ensure a degree of predictability and fairness? If the final decision remains unchanged despite a wide variation in the prior assumptions the verdict demonstrates a high level of robustness. This approach allows appellate courts to scrutinize the logical foundations of the decision, rather than second-guessing the hidden personal belief of the trier of fact.

5.1. The "Washing out" of the Prior

The focus of a rational trial should shift from *What is the exact prior?* to *Does the variation of the prior actually change the final decision?* At first, from a strictly subjectivist perspective, searching for an exact or true (prior) probability is a conceptual category mistake, as probability does not exist in nature but resides solely in the mind of the evaluator ("Probability is a state of mind not a state of nature" reported by Savage, 1972). Consequently, the task of the judge is not to estimate an objective frequency, but to assign a coherent value that transparently reflects their state of knowledge at a specific historical moment. Secondly, this approach relies on the principle that as the value of evidence (through its likelihood ratio value) increases, the influence of the prior odds on the posterior odds diminishes.

In many cases, the forensic evidence is so compelling (e.g., a high likelihood ratio value from DNA genotypes in paternity cases) that the posterior odds will point towards the same conclusion regardless of whether the initial prior was 1:2, 1:1,000 or 1:1,000,000. Through sensitivity analysis, we can demonstrate that the impact of the subjectivity of the judge is effectively "washed out" by the strength of the scientific findings (as long reported and proved in the statistical, judicial and forensic literature, see, e.g., Box and Tiao, 1973; Robertson *et al.* 2016). This provides a robust safeguard: the verdict is not just a consequence of the judge's initial bias, but a result of the evidence's dominance.

5.2. Transparency as a shield against arbitrariness

The true danger in legal fact-finding is not subjectivity itself, but *unstated* subjectivity. Judicial intuition is traditionally a black box that cannot be challenged. In contrast, an analysis of the effect of the choice of prior makes the judicial reasoning process auditable and transparent.

If a judge's conclusion is shown to be stable across a wide range of priors, the verdict gains a robustness that statistics based on frequency data can rarely provide in criminal cases which are generally unique by nature. This transparency allows for a more rigorous appellate review: the parties can debate the reasonableness of the prior range and the robustness of the evidence, rather than guessing the hidden thoughts of the trier of fact. Thus, analysis of the effect of the choice of prior transforms the prior challenge into an opportunity for greater judicial accountability.

6 THE DANGER OF THE CONTEXTUAL VACUUM AND THE PREVENTION OF MISCARRIAGES OF JUSTICE

In the final part of her article, Mackor (2026) reflects on the nature of Bayesian modelling as a tool for "Preventing miscarriages of justice" (The NWO research project, Mackor, 2026 at p. 20). While she expresses caution regarding the potential for misuse, we argue that the real danger to justice does not lie in the application of Bayesian reasoning, but in the absence of its application. Specifically, the greatest risk of error arises when evidence is evaluated in a contextual vacuum.

6.1. Context as a logical necessity

Evaluating evidence in a contextual vacuum is not merely a statistical limitation, it is a fundamental distortion of the legal truth. A piece of evidence does not possess an intrinsic meaning; its relevance is entirely dependent on the background information and the competing hypotheses. Moreover, facts in a courtroom do not exist in isolation, they are elements of a narrative. The Bayesian paradigm acts as a safeguard against this distortion, as it forces the decision-maker to formally acknowledge the impact of background information on the probability of the case hypotheses.

6.2. Misuse vs. invalidity

We must address the criticism that Bayesian ideas are inappropriate because it is very easy for them to be misused.The fact that a paradigm can be misused does not render the paradigm itself invalid. On the contrary, the Bayesian line of reasoning is

the only one that offers a coherent language and tool set to detect and correct such misuses. A model that makes its assumptions (priors) explicit is far easier to challenge and correct than a traditional holistic judgment where the reasoning remains hidden.

6.3. Bridging the gap: from data to justice

The reconciliation of subjective beliefs and frequency data is not just a statistical exercise, it is a requirement for fairness. A justice system that demands absolute objectivity where it cannot exist, such as in the unique circumstances of a criminal act, is a system that invites hidden biases. By embracing the subjectivist approach, the legal system moves toward a more honest form of rationality, where the uncertainty is quantified rather than ignored.

7. CONCLUSION: EMBRACING COHERENCE IN LEGAL FACT-FINDING

Anne Ruth Mackor's reflection on Dutch case law provides a timely opportunity to re-evaluate the role of Bayesian modelling in the courtroom. While the prior challenge is often cited as an insurmountable obstacle, we have demonstrated that this challenge is largely a product of an outdated frequentist view of probability.

By adopting a subjectivist perspective, we can conclude that:

1) Subjectivity is an asset, not a flaw, because it provides the necessary transparency for judicial reasoning.

2) The division of labour is key: forensic scientists evaluate the evidence through a coherent measure, the likelihood ratio, while judges provide the necessary context to assign the prior probabilities to the set of legally relevant hypotheses of interest to the court.

3) Robustness can be checked with sensitivity analysis. Such an analysis investigates how variations in prior specifications affect the outcome. This investigation ensures the legal process remains predictable and fair, even in the face of uncertainty. If the sensitivity analysis indicates a lack of robustness this casts doubt on the strength of the associated case.

The subjectivist approach does not make the inherent difficulties of criminal cases vanish, but it provides the only coherent framework for handling them. While we acknowledge that the debate on judicial objectivity is far from settled, we argue that a radical subjectivist approach offers a more logically consistent framework for the transparency of the trial. The Bayesian model is not merely a method to be added to the judge's toolkit, it is the logical foundation upon which a modern, transparent, and rational system of evidence evaluation and case interpretation must be built.

ACKNOWLEDGEMENTS

The authors thank the Swiss National Science Foundation (SNSF) for its support through Advanced Grant number TMAG-1 225782/1 (Forensic inference and aids to decision-making for scientific evidence).

REFERENCES

Aitken C.G.G., Gammerman A., Probabilistic reasoning in evidential assessment. *Journal of the Forensic Science Society* 29 (1989) 303-316.

Aitken C.G.G., Connolly T., Gammerman A., Zhang G., Bailey D., Gordon R., Oldfield R., Statistical modelling in specific case analysis. *Science & Justice* 36 (1996) 245-255.

Aitken C.G.G., Taroni F., Bozza S., Statistics and the evaluation of evidence for forensic scientists. 3rd ed., John Wiley & Sons, Chichester (2021).

Berger C.E.H., Finally a really forensic worldwide standard: ISO 21043 Forensic sciences, Part 4. Interpretation. Forensic Science International 10 (2025) 100589. *https://doi.org/10.1016/j.fsisyn.2025.100589*.

Bertillon A., De l'identification par les signalements anthropométriques, Imprimé à part des *Archives d'anthropologie criminelle et des sciences pénales*. G. Masson, Paris (1886) 17-18.

Biedermann A., Taroni F., Garbolino P., Equal prior probabilities: can one do any better? *Forensic Science International* 172 (2007) 85-93.

Biedermann A., Garbolino P., Taroni, F. (2013). The subjectivist interpretation of probability and the problem of individualization in forensic science. *Science & Justice* 53 (2013) 192-200.

Biedermann A., Bozza S., Taroni F., Aitken C.G.G., Reframing the debate: a question of probability, not of likelihood ratio. *Science & Justice* 56 (2016) 392-396.

Box G.E.P., Tiao G.C., *Bayesian inference in statistical analysis*. John Wiley & Sons, New York (1973).

Dawid A.P., Probability, causality and the empirical world: a Bayes - de Finetti - Popper - Borel synthesis. *Statistical Science* 19 (2004) 44-57.

Dawid A.P., Galavotti M.C., de Finetti's subjectivism, objective probability, and the empirical validation of probability assessment. In M.C. Galavotti (Ed.), *Bruno de Finetti, Radical probabilist*. College Publications, London (2009).

de Finetti B., Probability: the subjective approach. In R. Klibansky (Ed.), *La Philosophie Contemporaine*, vol. 2 (1968) 45-53. La Nuova Italia, Firenze.

de Finetti B., *Probability, induction and statistics - The art of guessing*. John Wiley & Sons, New York (1972).

de Finetti B., Probability does not exist. In Kyburg, H. E., & Smokler, H. E. (Eds), *Studies in Subjective Probability*. Wiley & Sons, New York (1974) 213-224.

de Finetti B., The proper approach to probability. In G. Koch, F. Spizzichino (Eds), Exchangeability in probability and statistics. North-Holland Publishing Company, Amsterdam (1982) 1-6.

Edwards W., Influence diagrams, Bayesian imperialism, and the Collins case: an appeal to reason. *Cardozo Law Review* 13 (1991) 1025-1074.

Eggleston R., Evidence, proof and probability. 2nd edn., Weidenfeld & Nicolson, London (1983).

ENFSI (2015). *ENFSI Guideline for evaluative reporting in forensic science*. Dublin, available at *http://enfsi.eu/documents/forensic-guidelines/*.

Fischhoff B., Beyth-Marom R., Hypothesis evaluation from a Bayesian perspective. *Psychological Review* 90 (1983) 239-260.

Kadane J.B., Prime time for Bayes. *Controlled Clinical Trials* 16 (1995) 313-318.

Kadane J.B., Schum D.A., *A probabilistic analysis of the Sacco and Vanzetti evidence*. John Wiley & Sons, New York (1996).

Kaye D.H., What is Bayesianism ? A guide for the perplexed. *Jurimetrics* 28 (1988) 161-177.

Lempert R. Modeling relevance. *Michigan Law Review* 75 (1977) 1021-1057.

Lindley D.V., Probabilities and the law. In D. Wendt C. Vlek (Eds), *Utility, probability, and human decision making*. D. Reidel Publishing Company, Dordrecht (1975) 223-232.

Lindley D., Probability. In C.G.G Aitken, D.A. Stoney (Eds), *The use of statistics in forensic science*. Ellis Horwood, Chichester (1991) 27-50.

Lindley D.V., *Understanding uncertainty*. Revised edn. JohnWiley & Sons, Hoboken (2014).

Locard E., *L'enquête criminelle. Traité de criminalistique*. Tome Septième. Livre VIII. Joannès Desvigne et Cie, Paris (1940).

Mackor A.R. Bayesian modelling of criminal cases as a whole – A philosophical reflection on Dutch case law. *Quaestio facti* (2026) DOI: 10.33115/udg_bib/qf.i10.23299

Press J., Tanur J.M., The subjectivity of scientists and the Bayesian approach. Dover Publications Inc., New York (2001).

Roberts P., Aitken C.G.G. *Practitioner guide 3: The logic of forensic proof: inferential reasoning in criminal evidence and forensic science* (2014), available at *https://rss.org.uk/news-publication/publications/law-guides/.RSS - Statistics law guides*

Roberston B., Vignaux G.A., Probability - The Logic of the Law. *Oxford Journal of Legal Studies* 13 (1993) 457-478.

Robertson B., Vignaux G.A., Berger C.E.H., *Interpreting evidence: evaluating forensic science in the courtroom*, 2nd ed., John Wiley & Sons, Chichester (2016).

Savage L.J., *The foundations of statistics*. 2nd revised edition, Dover Publications, Inc., New York (1972).

Schum D.A., *Evidential foundations of probabilistic reasoning*. John Wiley & Sons, Chichester (1994).

Taroni F., Biedermann A., Inadequacies of posterior probabilities for the assessment of scientific evidence. *Law, Probability and Risk* 4 (2005) 89-114.

Taroni F., Biedermann A., Bozza S., Garbolino P., Aitken C.G.G., *Bayesian networks for probabilistic inference and decision analysis in forensic science*. 2nd ed., John Wiley & Sons, Chichester (2014).

Taroni F., Garbolino P., Biedermann A., Aitken C.G.G., Bozza S., Reconciliation of subjective probabilities and frequencies in forensic science. *Law, Probability and Risk* 17 (2018) 243-264.

Taroni F., Garbolino P., Bozza S., Aitken C.G.G. The Bayes' factor: the coherent measure for hypothesis confirmation. *Law, Probability and Risk* 20 (2021) 15-36.

Tillers P., Introduction to Symposium on Probability and Inference in the Law of Evidence: the Uses and Limits of Bayesianism. *Boston University Law Review* 66 (1986) 381-390.

Tillers P. Green E.D. (Eds), *Probability and inference in the law of evidence – The uses and limits of Bayesianism*. Kluwer Academic Publishers, Dordrecht (1988).

Thompson W.C., Vuille J., Taroni F., Biedermann A., After uniqueness : the evolution of forensic science opinions. Judicature 102 (2018) 18-27.

Quaestio facti. Revista Internacional sobre Razonamiento Probatorio
Quaestio facti. International Journal on Evidential Legal Reasoning
Sección: Conjeturas y refutaciones
2026 I 11 pp. 157-180
Madrid, 2026
DOI: 10.33115/udg_bib/qf.i11.23278
Marcial Pons Ediciones Jurídicas y Sociales
© Marcello Di Bello
ISSN: 2604-6202
Recibido 5/5/26 | Aceptado 25/5/26 | Publicado *online*: __/__/2026
Editado bajo licencia Reconocimiento 4.0 Internacional de Creative Commons

EVENTS AND IDENTITY IN PROBABILISTIC MODELS OF LEGAL EVIDENCE

Marcello Di Bello
Arizona State University

ABSTRACT: This paper examines whether Bayesian networks are expressive enough to model reasoning with evidence in legal cases. Bayesian networks can represent many familiar patterns of evidential reasoning, including inferences from evidence to hypotheses, cumulative support from multiple items of evidence, and chains of inferences linking intermediate hypotheses to ultimate guilt. But other evidential inferences common in legal cases are more difficult to model. Focusing on a real criminal case, the paper distinguishes between identity- and event-level inferences. Event-level inferences show why certain actions amount to guilty conduct, while identity-level inferences link the defendant to those actions. The challenge for Bayesian network models of legal evidence is to represent how identity- and event-level inferences combine and reinforce one another. Meeting this challenge requires extending Bayesian networks beyond a purely propositional language.

KEYWORDS: bayesian networks; likelihood ratio; probability; criminal law.

1. INTRODUCTION

Can the evidence in a legal case be modeled probabilistically? Following Mackor (2026), this question is best divided into more specific subquestions. One concerns the target of the probabilistic modeling: are we asking whether individual items of evidence can be modeled probabilistically, or is the question about modeling the totality of the evidence in a case? Another subquestion concerns who is supposed to carry out the probabilistic analysis. Should the analysis be assigned to experts who

possess domain-specific knowledge, or should it remain with judges and jurors, perhaps with the assistance of court-appointed experts?

There is little doubt that individual items of evidence can be—and often are—evaluated probabilistically by experts. The most familiar examples come from forensic science (Taroni *et al.*, 2014). Suppose that genetic material matching the defendant is found at the crime scene. A forensic expert will typically offer the following probabilistic analysis of the match: they will assess how probable that finding would be if the defendant were the source of the genetic material, as compared with how probable it would be if someone else were the source. Expressed as a ratio, this comparison is known as the likelihood ratio:

$$\frac{\boldsymbol{P}\,(\textbf{ match} \mid \textbf{defendant is the source})}{\boldsymbol{P}\,(\textbf{ match} \mid \textbf{defendant is not the source})}$$

This ratio quantifies the degree of support that the match evidence provides for the source hypothesis: the greater the ratio, the stronger the support.

Forensic experts have also begun to apply probabilistic analyses beyond source-level propositions to activity-level propositions (Taylor *et al.*, 2018). The question of interest is often not merely whether the defendant is the source of a trace recovered at the crime scene, but how the trace came to be there, for instance, whether it got there during a violent confrontation between the defendant and the victim. How to carry out a probabilistic analysis of trace evidence for activity-level propositions is a topic of ongoing discussion in the literature (Stacey *et al.*, 2025).

A typical legal case, however, will consist of many pieces of evidence bearing on different propositions, not a single item of evidence bearing on a single proposition, such as a source- or activity-level hypothesis. So it is natural to ask whether the totality of the evidence in a legal case can be modeled probabilistically. That is a more complicated question. Which also makes the institutional question—who should carry out the probabilistic analysis?—more complicated. Experts may have the technical competence to build probabilistic models of individual items of evidence, but they lack the institutional authority to model a legal case as a whole. Judges and jurors, by contrast, have the authority to assess the totality of the evidence—that is what they are tasked with doing at trial—but they usually lack the technical competence required to construct a probabilistic model.

In this paper, I am going to set aside the institutional question. I will focus instead on whether it is in principle possible to model a case as a whole in probabilistic terms. A number of probabilistic models of entire legal cases have been developed. These models all rely on Bayesian networks (more on these soon). Examples include analyses of the Sacco and Vanzetti case (Kadane & Schum, 1996), the Anjum murders (Vlek *et al.*, 2014), the Simonshaven case (Fenton *et al.*, 2020), and a recent supermarket robbery case in the Netherlands (Hampson & Leeuwen, 2025). These examples suggest that probabilistic methods can be extended beyond individual items

of evidence to the analysis of a case as a whole using Bayesian networks. So there are some reasons for optimism.

But there are also reasons for skepticism. First, a common worry is that probabilistic models require numerical probabilities as inputs, yet the relevant data are often unavailable. Without data, the analysis must rely on rough estimates, educated guesses or expert judgment. But even when the data are available, they need not dictate the relevant probabilities. The data must first be interpreted and connected to the propositions at issue in the case. As Allen & Pardo (2007) note, the numbers used in probabilistic models are not "an objective datum that captures the value of evidence". Instead, they are "just more evidence, which itself must be interpreted to assess its value" (p. 119) [1].

A second challenge for probabilistic analyses of evidence, especially when the aim is to model an entire case, is the absence of a well-established methodology. Different analysts may produce different probabilistic models of the same case. To be sure, some authors have proposed ways to regiment the process of probabilistic model construction, for example, by identifying recurring idioms (Fenton *et al.*, 2013) or by comparing the modeling choices of different analysts (Hampson & Leeuwen, 2025). But the construction remains an interpretive task rather than a mechanical procedure. Any model—and probabilistic models are no exception—reflects choices about which propositions to include, at what level of abstraction to represent them, and how to structure the relations among them.

Now, even if the numerical and methodological challenges could be addressed, a third challenge would remain: are probabilistic models—and Bayesian networks in particular—expressive enough to capture the kinds of inferences that characterize reasoning with evidence in legal cases? This challenge is my focus here. As I will show, some evidential inferences can be represented quite naturally in Bayesian networks, while others are less tractable. To anticipate, the difficulty I will highlight lies in the interaction between inferences about events and inferences about identity. The integration of the two is not easy to model within standard (propositional) Bayesian networks. A richer formalism is needed.

The plan is as follows. I begin with a few examples of inferences that can be modeled with Bayesian networks (§2 and §3). To keep the discussion concrete, I turn to a simple legal case. I start with an informal description of the case and its items of

[1] A subjectivist Bayesian will respond that this is not a defect. Probabilities are not objective frequencies mechanically read off from the data. They are instead subjective degrees of belief, informed by data where available and revised in light of new evidence. On this view, probabilistic models do not remove judgment from the assessment of the evidence. Rather, they make judgment explicit and require it to be probabilistically coherent (Bozza *et al.*, 2026). Still, the probabilistic model does not, by itself, tell us whether the probability assignments it contains are well supported by the evidence. That further question depends on how the evidence is interpreted, how the propositions are framed, and whether the available data are relevant to the case at hand.

evidence (§4). I then present a coarse probabilistic model and refine it (§5 and §6). I show how some inferences resist probabilistic modeling in standard (propositional) Bayesian networks (§7 and §8). Finally, I conclude by underscoring the need for a more expressive formalism (§9).

2. BAYESIAN NETWORKS

I shall start with an overview of the central modeling tool used in probabilistic analyses of legal evidence: Bayesian networks. They are graphical representations of probabilistic dependencies among a set of variables. A Bayesian network has both a qualitative, graphical component and a numerical component[2]. Graphically, the network consists of nodes and directed edges (arrows) connecting the nodes. The graph must be acyclic (Figure 1): one cannot start from a node and, by following the direction of the arrows, eventually return to that same node.

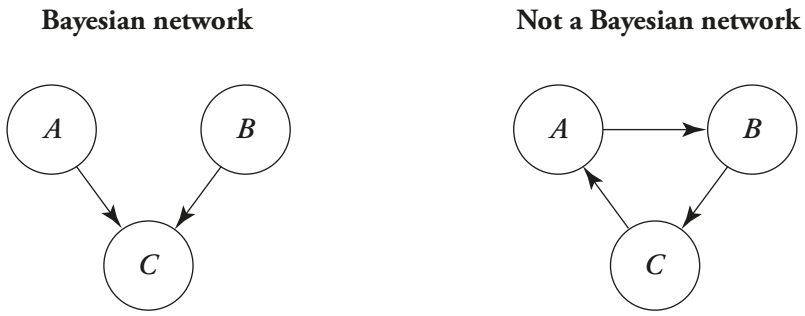
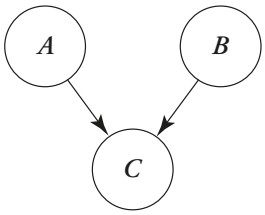


Figure 1. A Bayesian network must be directed and acyclic. The graph on the left satisfies both conditions. The graph on the right is directed, but not acyclic.

On the numerical side, each node in a Bayesian network represents a random variable that can take different possible values with different probabilities. Each node is associated with a probability table that specifies the probabilities of its possible values (Figure 2). In other words, the table specifies the probability distribution of the random variable. The directed edges indicate conditional probabilistic dependencies among the variables. If a node has no parents, its probability table will give a prior probability distribution. If a node has parents, its probability table will give a conditional probability distribution, specifying the probability of each possible value of that node given the values of its parents.

[2] For a book length introduction to Bayesian networks for modeling legal cases, see (Lagnado, 2021). For the application of Bayesian networks to forensic science, see (Taroni *et al.*, 2014).



$P(A = 1) = 0.3, \quad P(A = 1) = 0.7$

$P(B = 1) = 0.4, \quad P(B = 1) = 0.6$

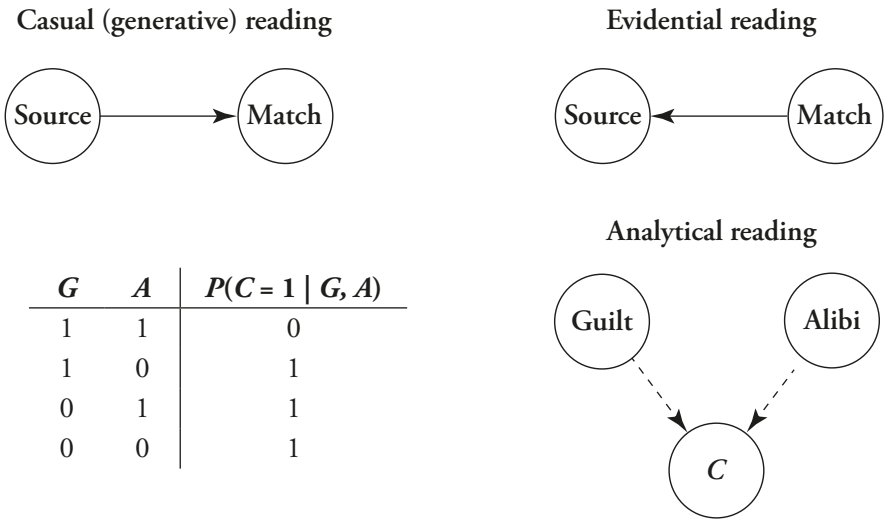
A	B	$P(C = 1 \mid A,B)$	$P(C = 0 \mid A,B)$
0	0	0.10	0.90
0	1	0.50	0.50
1	0	0.60	0.40
1	1	0.90	0.10

Figure 2. Nodes ***A*** and ***B*** have no parents and are assigned prior probabilities. Node ***C*** has parents ***A*** and ***B***, so it is assigned a conditional probability table.

A common source of confusion concerns the direction of the arrows. Strictly speaking, the difference between a network with an arrow from ***A*** to ***B*** and a network with an arrow from ***B*** to ***A*** is purely formal. In the first case, the model requires a conditional probability distribution for ***B*** given ***A***. In the second case, the model requires a conditional probability distribution for ***A*** given ***B***.

But arrows are often given an intuitive, less formal interpretation. On a common reading, an arrow from ***A*** to ***B*** suggests that ***A*** is a cause, or part of the causal explanation, of ***B***. For example, whether the defendant is the source causally influences whether the forensic test produces a match. So the graph will have an arrow from source to match. From an evidential point of view, however, the observed match is a reason for thinking that the defendant is the source. On that reading, the arrow will run from match to source (Dahlman & Kolflaath, 2022). In what follows, I will often use the causal reading. When evidence is modeled as the effect of an event, the arrow will run from the event to the evidence.

Arrows can also be given an analytical interpretation. For example, one node may be materially or conceptually incompatible with another: if one proposition is true, the other must be false. A node for the defendant's alibi and a node for the defendant's guilt are related in this way: if the defendant was elsewhere at the time of the offense, then the defendant could not have materially committed the offense. Or one node may be a consequence of another: if the defendant disposed of the weapon at a particular location, then the defendant must have been at that location. These relations reflect physical or conceptual constraints, and they can be represented in a Bayesian network by deterministic entries in the relevant conditional probability tables (Figure 3, bottom).



G	A	$P(C = 1 \mid G, A)$
1	1	0
1	0	1
0	1	1
0	0	1

Figure 3. In a causal reading, the arrow runs from cause to effect. In an evidential reading, reasoning proceeds in the opposite direction. The bottom panel illustrates an analytical constraint (note the dashed arrows): an alibi and guilt are materially incompatible. This can be represented by a constraint node ***C***, whose conditional probability table assigns probability zero to the combination of guilt and alibi.

Bayesian networks provide compact representations of joint probability distributions over many variables. The basic idea is simple: instead of writing out an enormous probability distribution that covers every possible combination of variables, the network breaks the problem into smaller pieces. Each variable is represented by a node, and the probability of that variable depends only on its parent variables, that is, the variables that point directly into it. Stated more formally, the joint probability distribution over all variables in the network can be factorized as a product of local conditional probabilities, each associated with a node and its parents. If the variables are $X_1, \ldots, X_n$, the joint distribution is written as the following product:

$$P(X_1, \ldots, X_n) = \prod_{i=1}^{n} P(X_i \mid \mathrm{Pa}(X_i))$$

where $\mathrm{Pa}(X_i)$ denotes the parents of X_i in the graph. The formula expresses the central assumption of a Bayesian network: once the parents of a variable are fixed, other non-descendant variables do not make any further difference to the probability of that variable. To put it more formally, the equality holds by the local Mark-

ov property of Bayesian networks: each node is conditionally independent of its non-descendants given its parents (Figure 4).

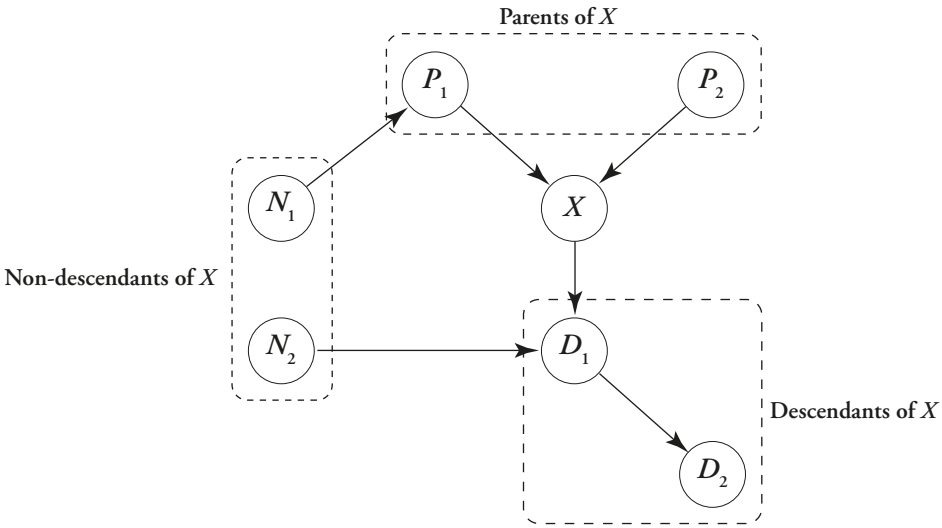


Figure 4. Once X's parents P_1 and P_2 are fixed, X's non-descendants N_1 and N_2 provide no further information about X. So, by the local Markov property, it follows that $P(X \mid P_1, P_2, N_1, N_2) = P(X \mid P_1, P_2)$.

3. IDIOMS IN BAYESIAN NETWORKS

It is now time to consider how Bayesian networks can model recurring patterns of inference in legal cases. In the diagrams that follow, some nodes are evidence nodes (represented by rectangles) while others are hypothesis nodes (represented by circles). A Bayesian network can thus be understood as a structure for representing probabilistic relationships or inferences between evidence and hypotheses.

In what follows, I will slightly abuse notation. I will use the same symbol both for the random variable represented by a node and for the proposition that this variable takes the value **1**. Thus, if S is a node representing the source hypothesis, then S denotes the random variable that can take the values $S = 1$ and $S = 0$. At the same time, I will use S as shorthand for the proposition that the source hypothesis is true (that is, $S = 1$) and $\neg S$ as shorthand for the proposition that it is false (that is, $S = 0$). Context should make clear which use is intended.

Following recurring inference patterns or idioms for Bayesian networks (Fenton *et al.*, 2013), I will start with the simplest one. The support that a piece of evidence lends to a hypothesis can be modeled by comparing the probability of observing the

evidence if the hypothesis is true with the probability of observing the same evidence if the hypothesis is false. As seen before, this is the likelihood ratio:

$$\frac{P(E \mid H)}{P(E \mid \neg H)}$$

If the ratio is greater than one, a rational decision-maker should increase their assessment of the probability of the hypothesis[3]. Graphically (figure 5, left side), this inference is modeled by the simplest Bayesian network consisting of just two nodes, one for the hypothesis and another for the evidence, with an arrow from hypothesis to evidence. Crucially, an arrow from H to E does not by itself mean that E supports H. That depends on the corresponding conditional probability table. If the table says that $P(E \mid H) > P(E \mid \neg H)$, then E supports H; if $P(E \mid H) = P(E \mid \neg H)$, then E is evidentially neutral; and if $P(E \mid H) < P(E \mid \neg H)$, then E counts against H. Thus, the graph signals a potential dependency. Whether that dependency amounts to support, neutrality, or opposition depends on the numbers in the probability table.

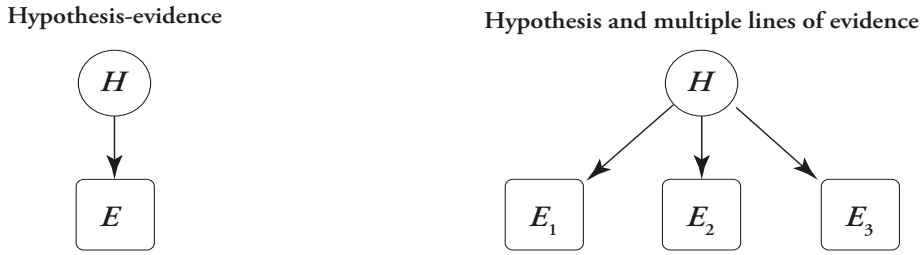


Figure 5. Simple inference patterns in Bayesian network form.

A more complex pattern arises when multiple items of evidence, $E_1, E_2, \ldots, E_k$, constitute independent lines of support for the same hypothesis H. This evidential pattern has a formal equivalent, that is, each E_i can be assumed to be conditionally independent given the hypothesis of interest (and given its negation)[4]. Then, their combined evidential force is modeled by multiplying their likelihood ratios:

---

[3] By the odds form of Bayes' theorem,

$$\frac{P(H \mid E)}{P(\neg H \mid E)} = \frac{P(E \mid H)}{P(E \mid \neg H)} \cdot \frac{P(H)}{P(\neg H)}$$

If the likelihood ratio is greater than one, then the posterior odds of H are greater than the prior odds of H. Provided $0 < P(H) < 1$, it follows that $P(H \mid E) > P(H)$, or in other words, the observation of the evidence rationally supports an increased degree of belief in the hypothesis.

[4] In symbols, for three items of evidence E_1, E_2, E_3:

$P(E_1,E_2,E_3|H) = P(E_1|H)P(E_2|H)P(E_3|H)$,

$P(E_1,E_2,E_3|\neg H) = P(E_1|\neg H)P(E_2|\neg H)P(E_3|\neg H)$

$$\frac{P(E_1|H)}{P(E_1|\neg H)} \times \frac{P(E_2|H)}{P(E_2|\neg H)} \times \ldots \times \frac{P(E_k|H)}{P(E_k|\neg H)}$$

If each likelihood ratio is greater than one, a rational decision-maker should increase their assessment of the probability of the hypothesis by a significant margin. Multiple, independent lines of evidence in support of ***H*** lend much stronger support taken together than each one in isolation. Graphically (figure 5, right side), this inference is modeled by a Bayesian network consisting of a hypothesis node from which outgoing arrows point to different evidence nodes.

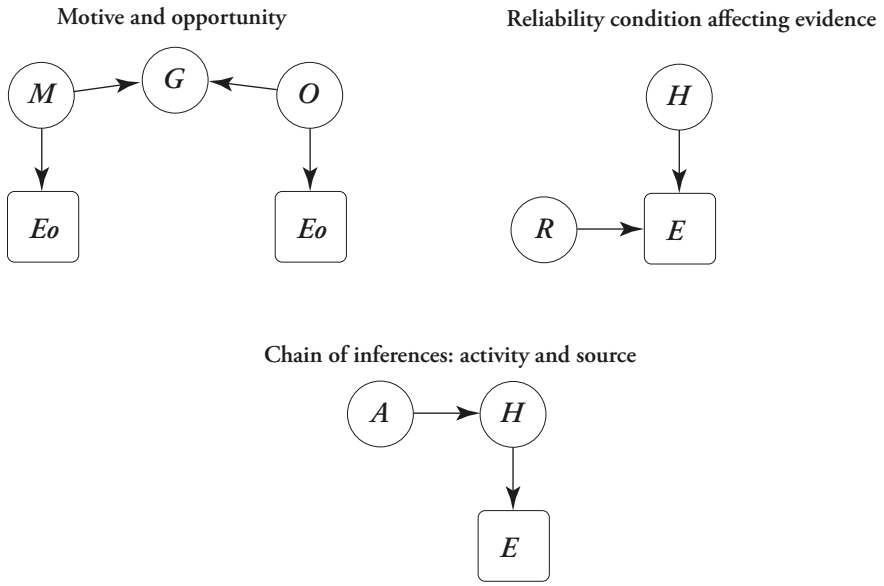


Figure 6. Other inference patterns in Bayesian network form.

Fenton *et al.* (2013) identify many other recurring inference patterns, including inferences from someone's motive and opportunity to their guilt. Having a motive and having the opportunity to commit a crime, for example by being near the crime scene, are both predictors of guilt. In a Bayesian network, this can be represented by nodes for motive and opportunity, each with an arrow pointing to the guilt node. Motive and opportunity may in turn be supported by evidence, so that they function as intermediate hypotheses within the network (Figure 6, top left side).

Another common pattern is the reliability (or accuracy) idiom. When a piece of evidence supports a hypothesis, its probative value depends on the reliability of the evidence. Suppose, for example, that an eyewitness states that the defendant was near the crime scene. That testimony supports the hypothesis that the defendant was near the scene only if the witness is reliable, say was not far away or did not make the

identification in the dark. Graphically (figure 6, top right side), this can be represented by a network in which both the evidence node and a reliability node bear on the hypothesis. In likelihood ratio terms:

$$\frac{P(E \mid H,R)}{P(E \mid \neg H,R)} > 1 \quad \frac{P(E \mid H,\neg R)}{P(E \mid \neg H,\neg R)} = 1$$

The first expression says that, conditional on the witness being reliable, the testimony is more probable if the hypothesis is true than if it is false. So the evidence supports the hypothesis. The second says that, conditional on the witness being unreliable, the testimony is equally probable whether the hypothesis is true or false. So the evidence provides no support when the reliability condition is not met.

Finally, consider the inference from match evidence to a source proposition, and from the source proposition to an activity proposition (Figure 6, bottom). The inference proceeds in two steps. First, the match evidence supports, to some degree, the source proposition, for example, that the trace recovered at the crime scene came from the defendant. Second, the source proposition supports, to some degree, an activity proposition, for example, that the defendant participated in the relevant criminal conduct. These two inferential steps can come apart. A DNA match could strongly support the hypothesis that the defendant was the source of the trace, while it remains unclear whether the trace was deposited during the crime or through innocent contact. Conversely, even if the source-to-activity inference is strong, the overall inference can still be weak if the matching evidence provides only limited support for the source proposition.

The inferential patterns identified above—the simple evidence-hypothesis idiom; multiple lines of evidence; motive and opportunity; reliability; chain of inferences—can also be combined in a modular fashion to form more complex models that consist of several items of evidence and hypotheses. Hampson & Leeuwen (2025) rely on those very idioms to model a supermarket robbery, and Fenton *et al.* (2020) do something similar with the Simonshaven case. This suggests that Bayesian networks have considerable expressive power: by combining common inferential idioms, they can model some legal cases in their entirety.

4. STATE VERSUS EVANS

Having seen some of the inferences that Bayesian networks can model, let us now turn to an actual legal case to keep the discussion concrete. The case is useful for our discussion because it is simple to understand but also complex enough to pose a challenge for probabilistic modeling. The prosecution's theory depends on multiple strands of evidence that support different propositions. Some items bear on motive, some on identity, some on presence near the crime scene, and others on possible post-offense conduct.

In *State v. Evans*, SC 21006 (Conn. Aug. 12, 2025), the prosecution alleged that the defendant, Evans, shot and killed Reginald May outside May's apartment in Bridgeport, Connecticut, shortly after 4:00 a.m. on July 2, 2017. The State's theory was that the killing was connected to an earlier dispute between Evans and the victim. Several items of evidence, taken together, link Evans to the shooting.

Evans and May had argued and fought over money during a moving job in Virginia on June 28, 2017. After that incident, May returned to Connecticut. A few days later, surveillance video from near May's apartment showed a large dark SUV circling the area and then parking on Alice Street at about 4:17 a.m. The footage showed a man in a hooded sweatshirt exiting the SUV, walking toward the apartment building, and returning to the vehicle at about 4:21 a.m.

The State sought to connect this footage to Evans in several ways. First, the SUV shown in the video was consistent with Evans's Cadillac Escalade. Second, police extracted a still image from the surveillance footage, and both the victim's brother, John May, and the victim's girlfriend, Cherry Williams, identified the person in the still image as Evans. The State also introduced evidence of a firearm recovered on July 13, 2017, at West Rock Nature Center. Ballistic examination linked the firearm to the crime. Finally, the State relied on phone location information showing that on July 3, 2017, Evans's phone connected to a cell tower near West Rock Ridge State Park, near the area where the firearm linked to the crime was found.

A list of the different items of evidence is provided below for ease of reference:

Surveillance video. Surveillance footage on July 2, 2017 from near the victim's apartment showed a dark SUV parking on Alice Street around 4:17 a.m. A man in a hooded sweatshirt exited the SUV, walked toward the building, and returned to the SUV around 4:21 a.m.

ID from video still. A still image extracted from the surveillance footage was shown to John May and Cherry Williams, both of whom identified the person in the image as Evans.

SUV evidence. SUV in the video was consistent with Evans's Cadillac Escalade.

Firearm evidence. A firearm and extended magazine were recovered on July 13, 2017, at West Rock Nature Center, followed by ballistic examination linking the firearm to the shooting.

Phone location. On July 3, 2017, Evans's phone connected to a cell tower near West Rock Ridge State Park, close to the location where the firearm connected to the crime was recovered.

5. INDEPENDENT LINES OF EVIDENCE

As a first attempt, we can model this case with a Bayesian network that uses the five key pieces of evidence all at once: ID, SUV, firearm, phone and surveillance vid-

eo evidence. Each of them links the defendant to the crime in its own way: via the phone location and recovery of the firearm in that location; similarity of the SUV in the surveillance video with defendant's vehicle; and identification of the person in the video. The structure in Figure 7 illustrates the textbook case for aggregating independent items of evidence, all bearing on the same hypothesis.

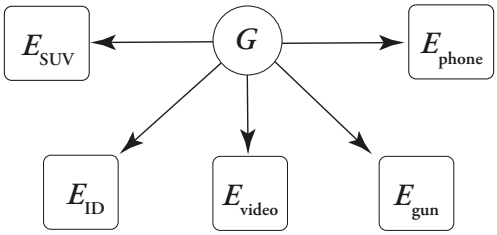


Figure 7. Simplest Bayesian network: ID, SUV, phone, firearm and video evidence are treated as conditionally independent indicators bearing on guilt.

Quantitatively, each piece of evidence can be paired with a measure of how strongly, on its own, it supports the guilt hypothesis, such as the likelihood ratio:

$$LR(E_i;G) = \frac{P(E_i \mid G)}{P(E_i \mid \neg G)}$$

In this model, the combined strength of support of the five pieces of evidence in favor of the guilt hypothesis is the product of the individual likelihood ratios:

$$LR(E_{ID};\, G) \cdot LR(E_{SUV};\, G) \cdot LR(E_{PHONE};\, G) \cdot LR(E_{GUN};\, G) \cdot LR(E_{VIDEO};\, G)$$

Suppose that the ID, SUV, phone, firearm and video evidence have likelihood ratios of 3, 5, 4, 2 and 2. The numbers are illustrative. Then, their combined likelihood ratio will be **3×5×4×2×2 = 240**. To have a sense of the strength of such evidence, in the aggregate, a likelihood ratio of 240 would be able to turn a 1% initial probability of guilt (1-to-99 prior odds) into approximately 70% posterior probability of guilt (2.4-to-1 posterior odds)[5].

But there is a complication. Our model treats the evidence as a collection of independent items all bearing on the guilt hypothesis, as if each were a standalone indicator of guilt. In fact, the different pieces of evidence speak to different hypotheses. The phone evidence links the defendant to the crime by showing that the perpetrator disposed of the weapon in the same area that the defendant visited. The video, ID and SUV evidence place the defendant outside a building at a certain time. The video

[5] The calculation uses the odds form of Bayes' theorem: posterior odds = prior odds × likelihood ratio. A prior probability of **1%** corresponds to prior odds **1/99**. Multiplying by **240** gives posterior odds **240/99≈2.42**. Converting odds back into probability gives **2.42/(1+2.42)≈71%**.

shows what the person outside the building did; the ID and SUV evidence connect the person in the video to the defendant. Thus, the model of multiple independent lines of evidence, each supporting guilt, is too simplistic.

6. OPPORTUNITY → GUILT → DISPOSAL

To refine our model, I shall now distinguish between the opportunity hypothesis and the guilt hypothesis, and between the guilt hypothesis and disposal hypothesis. The opportunity hypothesis (OPP) is the claim that Evans was present near the crime scene around the time of the shooting. The guilt hypothesis (G) is the stronger claim that Evans committed the shooting. The surveillance video ID and the SUV evidence are naturally interpreted as evidence that Evans was in the relevant location at the relevant time and thus had the opportunity to commit the crime. Evans' guilt is supported through the further inference from OPP to G. We can represent this structure as a Bayesian network (Figure 8, right side) in which both evidentiary nodes support OPP, and OPP in turn supports G.

Next, we will add two other evidence nodes: the recovery of the firearm and the post crime phone location evidence. The two items of evidence support—how, exactly, needs to be seen—the intermediate proposition that Evans disposed of the weapon. So we can add a disposal node to the network, call it ***D***. We can complete the model by drawing arrows from the disposal node to the evidence nodes, one for evidence about where the weapon was found and another for the phone location. The disposal hypothesis supports guilt via further inference.

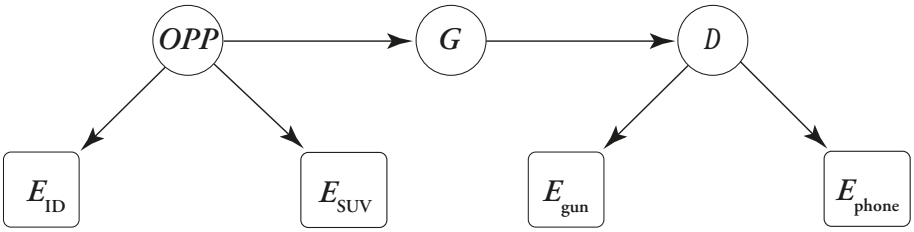


Figure 8. Bayesian network distinguishing guilt (***G***) from post-crime disposal of the firearm (***D***) and opportunity (***OPP***). The evidence of the recovered firearm and the phone-location evidence most naturally bear on disposal, which in turn supports guilt. The ID and SUV evidence, instead, most naturally bear on opportunity.

The new model (Figure 8) is more faithful to the evidential structure of the case than a flatter model in which each evidential item supports guilt. The new model makes explicit that the evidence does not bear on guilt in a simple one-step way. Instead, the evidence supports intermediate events, and these in turn support guilt. We have two independent evidential routes, one that goes through opportunity, and

another that goes through the disposal hypothesis. The contributions of each route add up. But how, exactly? The math here is more involved than before.

Let us start with the opportunity route. In the model, there are two independent lines of evidence bearing on the opportunity hypothesis, E_{ID} and E_{SUV}. (I'll revisit this modeling assumption later.) By conditional independence, the likelihood ratio for the opportunity hypothesis is

$$LR(E_{ID}, E_{SUV};OPP)= \frac{P(E_{ID}|OPP)}{P(E_{ID}|\neg OPP)} \cdot \frac{P(E_{SUV}|OPP)}{P(E_{SVU}|\neg OPP)}$$

The problem now is that we need to make a further inference from opportunity to guilt. The likelihood ratio relative to opportunity, based on the ID and SUV evidence, must be transferred to the guilt hypothesis. This is the formula we need[6]:

$$LR(E_{ID}, E_{SUV};OPP)= \frac{LR(E_{ID}, E_{SUV};OPP)\ P(OPP|\ G)+1\ P(OPP|\ G)}{LR(E_{ID}, E_{SUV};OPP)\ P(OPP|\ \neg G)+1\ P(OPP|\ \neg G)}$$

We can run a similar argument for the evidence supporting the disposal hypothesis which in turn supports the guilt hypothesis. This is the other evidential route. In the model, there are two items of evidence bearing on the disposal hypothesis, E_{gun} and E_{phone}. Again by conditional independence, the combined likelihood ratio for the disposal hypothesis is

$$LR(E_{gun}, E_{phone};D)= \frac{P(E_{gun}|D)}{P(E_{gun}|\neg D)} \cdot \frac{P(E_{phone}|D)}{P(E_{phone}|\neg D)}$$

As before, the likelihood ratio relative to the disposal hypothesis, based on gun and phone evidence, must also be transferred to the guilt hypothesis:

$$LR(E_{gun}, E_{phone};G)= \frac{LR(E_{gun}, E_{phone};D)\ P(D|G)\ +1\ -\ P(D|G)}{LR(E_{gun}, E_{phone};D)\ P(D|\neg G)\ +1\ -\ P(D|\neg G)}$$

So, evidence that more closely supports ***OPP*** must be translated into support for ***G*** through the opportunity-to-guilt route; evidence that more closely supports ***D*** must also be translated into support for ***G*** through the disposal-to-guilt route. As a rule of thumb, the inference toward guilt along each route is constrained by

[6] Let $E = (E_{ID}, E_{SUV})$.Assume that E bears on guilt G only through the opportunity proposition ***OPP***. Then, by the law of total probability,

$$\frac{P(E\,|G)}{P(E\,|\neg G)} = \frac{P(E\,|\,OPP)\ P(OPP\,|\,G) + P(E\,|\,\neg OPP)\ P(\neg OPP\,|\,G)}{P(E\,|\,OPP)\ P(OPP\,|\,\neg G) + P(E\,|\,\neg OPP)\ P(\neg OPP\,|\,\neg G)}$$

Dividing numerator and denominator by $P(E\,|\,\neg OPP)$ we obtain the formula.

the weakest inferential link. In other words, the evidential force transmitted to ***G*** is bounded both by how strongly the evidence supports ***OPP*** and by how strongly ***OPP*** supports ***G***. A weak link at either stage acts as a bottleneck. More formally,

$$LR(E_{ID}, E_{SUV}; G) \leq \min \left\{ LR(E_{ID}, E_{SUV}; OPP), \frac{P(OPP \mid G)}{P(OPP \mid \neg G} \right\}$$

Similarly, along the route from disposal to guilt, the inference is constrained by the weakest inferential link:

$$LR(E_{gun}, E_{phone}; G) \leq \min \left\{ LR(E_{gun}, E_{phone}; D), \frac{P(D \mid G)}{P(D \mid \neg G} \right\}$$

Finally, the overall likelihood ratio for the guilt hypothesis, based on all the evidence we have, is the result of the following multiplication:

$$LR(E_{ID}, E_{SUV,} E_{gun}, E_{phone}; G) = LR(E_{ID}, E_{SUV,}; G) \cdot LR(E_{gun}, E_{phone}; G)$$

This holds because the items of evidence along each evidential route are conditionally independent given the guilt hypothesis, or in other words, the opportunity-related evidence and disposal-related evidence each provide an independent route of support for guilt. So the support for ***G*** adds up across the two evidential routes and the relevant likelihood ratios multiply. But since each route is bounded by the weakest inferential link, when the two routes are aggregated, the overall strength of the inference is bounded by the product of the weakest links in each route. All in all, the total evidential force accumulate across independent evidential routes, but only after each route has been limited by its weakest inferential link.

7. EVENT AND IDENTIFICATION INFERENCES

The model in the previous section, albeit an improvement on the earlier one, suffers from a number of problems. It incorporates four sources of evidence: SUV, ID, phone location and firearm evidence. But it leaves out a fifth, namely the content of the surveillance video. The video shows a dark SUV arriving near the apartment complex shortly before the shooting, a person exiting the SUV and walking toward the apartment complex, and that person returning shortly afterward. From the sequence, location and timing of events, we can infer that the person who engaged in the actions depicted in the video almost certainly committed the shooting. But, from a modeling standpoint, it is not clear where the video evidence should be placed in the inferential structure represented by the network. It does not bear in any natural way on the opportunity hypothesis, at least not if opportunity is understood as the proposition that Evans was near the crime scene. The video evidence, by itself, does not link Evans to the crime. It only links the person in the video to the shooter, but leaves open who that person may be.

These observations also suggests that the model misrepresented the inferential role of the ID and SUV evidence. The model takes those items to be evidence for the opportunity hypothesis that Evans was near the crime scene. But, ID and SUV evidence do not independently support that hypothesis. Rather, they help identify the person in the video as Evans and the vehicle as his SUV. Only when that identification is combined with the content of the video can we infer that Evans was near the crime scene at the relevant time.

A third problem—related to the second—is that the proposed model misleadingly suggests that phone and firearm evidence, separately, support the disposal hypothesis. In fact, one without the other would be irrelevant. The recovery of the firearm at the remote location does not by itself show who placed it there. Nor does the suspect's phone being located near that site, by itself, establish involvement in the disposal of the firearm. Each item seems to derive much of its significance from the other. The firearm evidence makes the phone evidence probative because the suspect appears near the place where the crime weapon was found. The phone evidence makes the firearm evidence probative because the suspect appears at the place where the weapon was apparently discarded. The force of the evidence therefore arise from their fit within a narrative of post-crime disposal.

One way to diagnose what is going on is to distinguish two distinct inferential tasks, and possibly more. The first task is event reasoning: what happened, how, where and when? Here the aim is to determine the sequence of events in space and time. The video and firearm evidence address event reasoning. Someone approached the location, as shown in the video; the victim was killed; the firearm used in the crime was later disposed of in a remote location; and the recovered firearm is linked to the crime by forensic analysis. The second task is identity reasoning: who did it? Once the event sequence is fixed, the question becomes whether the defendant is the person who carried out those actions. The SUV and ID evidence address the identity question prior to the crime: who was the person around the crime scene? The phone location evidence, instead, addresses the identity question post-crime, specifically whether the defendant was the person who went to the disposal site[7].

If this analysis is right, a more adequate model should clearly distinguish hypotheses about events from hypotheses about identity, and then combine the two. One way to proceed is to build two preliminary submodels (Figure 9). The event model will contain propositions such as: an individual approached the scene, killed the victim, and later disposed of the firearm at the remote location. The identity model would then ask whether that individual was the defendant. The central challenge is then to explain how these two submodels can be combined.

[7] Tribe (1971) distinguishes between (1) evidence "directed to the occurrence or nonoccurrence of the event, act, or type of conduct on which the litigation is premised" and (2) evidence "directed to the identity of the individual responsible for a certain act or set of acts." (p. 1339)

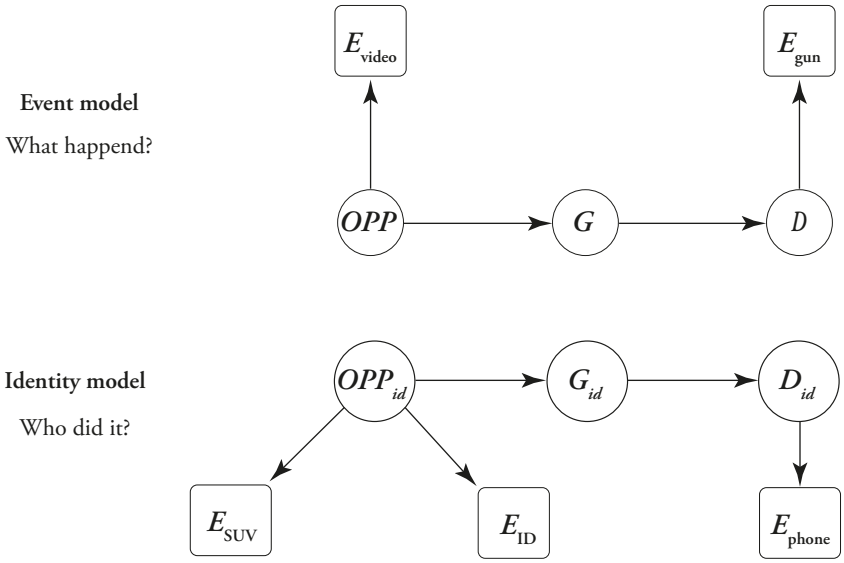


Figure 9. Two Bayesian networks with the same structure, $\boldsymbol{OPP \rightarrow G \rightarrow D}$, but different interpretations. In the event model, the nodes represent propositions about the sequence of events: opportunity, commission of the crime, and later disposal of the firearm. In the identity model, the same nodes represent propositions about whether the defendant is the relevant actor at each stage.

8. COMBINING THE TWO

In an attempt to build a unified model that does not confuse identity and event inferences, I will focus on three key nodes: opportunity, disposal and guilt.

8.1. Opportunity

To fix some notation, let $\boldsymbol{E_{ID}}$ and $\boldsymbol{E_{SUV}}$ be the identification evidence that supports the proposition ***ID*** that the defendant is the person depicted in the video. Let $\boldsymbol{E_{video}}$ be the evidence that supports the proposition ***N(v)*** that the person depicted in the video was near the crime scene at the relevant time. This way the opportunity hypothesis has been disambiguated into two distinct hypotheses, one about the events (***N(v)***) and another about identity (***ID***). But the hypothesis we care about is ***N(d)***: that the defendant was near the crime scene. How should this be modeled?

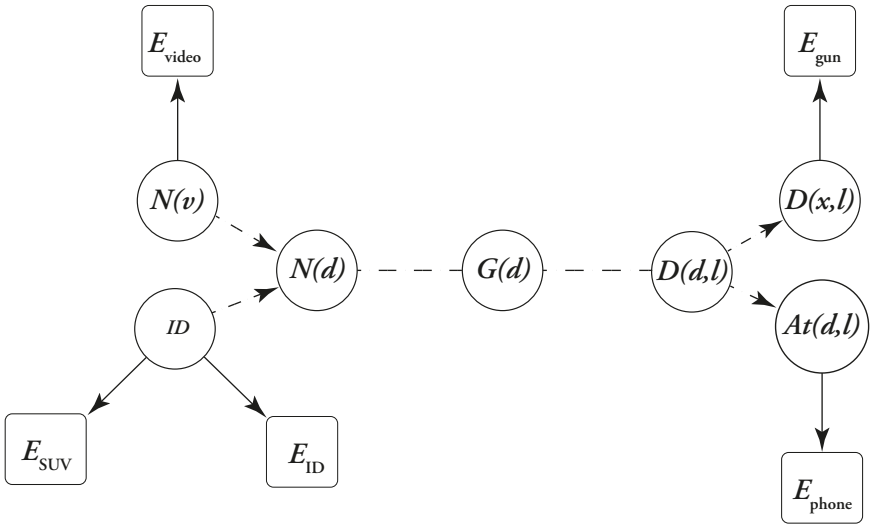


Figure 10. (Left) ***ID*** *is the proposition that the defendant is the person depicted in the video;* ***N(v)*** *is the proposition that the person depicted in the video was near the crime scene. The opportunity proposition* ***N(d)*** *is modeled as depending on* ***ID*** *and* ***Nv****. (Right)* ***D(x, l)*** *is the proposition that a person* ***x*** *disposed of the firearm at location* ***l****;* ***At(d, l)*** *is the proposition that the defendant* ***d*** *was at location* ***l****. The dashed arrows indicate analytical dependencies. The dashed lines indicate unspecified dependencies.*

The simplest approach is to think of it as a conjunction:

$$N(d) \coloneqq ID \wedge N(v)$$

But this reading cannot be right. Although the conjunction of the two propositions on the right entails that the defendant was near the crime scene, the converse does not follow. If the defendant was near the crime scene, it does not necessarily follow that he was the person in the video. The defendant might have been near the crime scene in some other way, independently of the events depicted in the surveillance footage. In addition, if ***N(d)*** were a mere conjunction of the other two hypotheses, each item of evidence would behave as though it could make an independent contribution toward supporting ***N(d)***[8]. But, on its own, neither item is probative of

[8] Assume that E_{ID} supports ID, so that
$P(ID \mid E_{ID}) > P(ID)$
Does E_{ID} also support $N(d)$? Since
$P(N(d) \mid E_{ID}) = P(ID \wedge N(v) \mid E_{ID}) = P(ID \mid E_{ID})\, P(N(v) \mid ID)$
By definition
$P(N(d)) = P(ID)\, P(N(v) \mid ID)$
The question is whether
$P(ID \mid E_{ID})\, P(N(v) \mid ID) > P(ID)\, P(N(v) \mid ID)$
The inequality holds assuming $P(ID \mid E_{ID}) > P(ID)$ unless $P(N(v) \mid ID) = 0$.

the defendant's presence near the scene. As noted before, their evidential force arises only when they are considered together[9].

Here is a better interpretation of the opportunity hypothesis and the inference that supports it. The video evidence supports the proposition that the person shown in the video was near the crime scene (***N(v)***), while the identification evidence supports the proposition that the person shown in the video is the defendant. To make fully transparent the identification between the person ***v*** in the video and the defendant ***d***, let us denote this proposition by ***d = v*** instead of the more generic ***ID***. The inference in question that leads to ***N(d)*** looks like this:

$$\frac{N(v) \qquad d=v}{N(d)}$$

The inference is deductive: once ***N(v)*** and ***d = v*** are accepted, ***N(d)*** follows by substituting co-referring expressions. How does this pattern of inference fit into a standard Bayesian network?

One possible response is to hard code the inference into the conditional probability table for node ***N(d)*** representing the hypothesis that the defendant was near the crime scene. The node will have two parents: ***N(v)*** and ***v=d*** or ***ID*** (Figure 10, left side). The probability table could then be specified as follows:

$N(v)$	$v = d$	$P(N(d) = 1 \mid N(v), v = d)$
1	1	1
0	1	q
1	0	p
0	0	p

The first row is key. It captures the deductive inference we are after: if the person in the video was near the crime scene, and if that person was the defendant, then the defendant was near the crime scene. In that case, ***N(d)*** follows with probability **1**. The remaining rows assign baseline probabilities ***p*** or ***q***[10].

[9] This problem can perhaps be deflected by noting that the evidential support will often be limited. Suppose $P(ID) = 1/1\,000\,000 = 10^{-6}$ since 1 million people (or whatever the number) could have been in the video, and similarly $P(N(v)) = 10^{-6}$ since the video could depict events in many possible locations. Assuming probabilistic independence, $P(ID \wedge N(v)) = 10^{-12}$. If the identification evidence E_{ID} makes ***ID*** almost certain, say $P(ID \mid E_{ID}) = .99$, then

$P(N(v) \wedge ID \mid E_{ID}) \approx P(ID \mid E_{ID})\, P(N(v)) = .99 \times 10^{-6} \approx 10^{-6}$

The probability increase is significant in relative term, but still tiny in absolute. A similar point applies for the video evidence and proposition (***N(v)***).

[10] The second and third rows reflect the idea that, if the person in the video was not the defendant, then facts about the person in the video do not by themselves support or oppose the proposition that the defendant was near the crime scene.

This construction shows that the substitution inference can be represented in a Bayesian network by the appropriate assignment of values in the conditional probability table. The limitation, however, is that the model does not fully express the pattern of inference involved. The underlying reasoning relies on coreference: from ***N(x)*** and ***x = d***, infer ***N(d)***, for any ***x***. A standard (propositional) Bayesian network can encode an instance of the pattern of inference, not the general pattern.

8.2. Disposal

Another modeling challenge arises for the disposal hypothesis. To start, the firearm evidence $\boldsymbol{E}_{\mathbf{gun}}$ supports the proposition that someone disposed of a firearm at a definite location. Ballistic evidence—which we have not separately modeled—supports the claim that the firearm recovered at that location was the firearm used in the shooting. So, the firearm evidence support the claim—call it ***D(x,l)*** —that someone disposed of the firearm at that specific location. Who did it? The phone location evidence $\boldsymbol{E}_{\mathbf{phone}}$ supports the proposition that the defendant was at that same location, call it ***At(d,l)***. Given that the location is the same, these propositions—taken together, not independently—support the conclusion ***D(d,l)*** that it was the defendant who disposed of the firearm at that location.

As with the inference about opportunity, the conclusion ***D(d,l)*** that the defendant disposed of the firearm at that location cannot be modeled as a conjunction of the other two propositions ***D(x,l)*** and ***At(d,l)***. The inference about the disposal hypothesis can be formalized, as follows:

$$\frac{\exists x D(x, l) \quad At(d, l) \quad \neg Defeater}{D(d, l)}$$

The inference is defeasible because the defendant's presence at ***l*** does not deductively establish that he was the person who disposed of the firearm there, although the identity of the locations makes it somewhat likely.

How does this pattern of inference fit into a standard Bayesian network? The proposition ***D(d, l)*** that the defendant disposed of the firearm at location ***l*** is an explanatory hypothesis. If true, it would make sense of the truth of two otherwise separate intermediate hypotheses: first, that the firearm was disposed of at location ***l***, and second, that the defendant was at that location. So the most natural option is to assign a node for hypothesis ***D(d, l)*** with two outgoing arrows pointing to two nodes, one for ***∃xD(x, l)*** and another for ***At(d, l)*** (Figure 10, right side). This representation is natural since those hypotheses are two consequences of ***D(d, l)***. The following conditional probability tables capture the relationships:

$D(d, l)$	$P(At(d, l) = 1 \mid D(d, l))$
1	1
0	p

$D(d, l)$	$P(d(x, l) = 1 \mid D(d, l))$
1	1
0	q

The first row of each table captures the relevant implication. The second row in each table is a baseline value. If the defendant did not dispose of the weapon at the location, it does not follow that he was not there or that no one disposed of the weapon there. The defendant may have been there for some other reason or someone else may have disposed of the weapon.

This representation does the trick, but again it has a limitation. The probability tables do not explicitly encode the preservation of the same location ***l*** across propositions. As with the video and identification inference, the network can hard code the inference, but does not fully capture the general pattern of reasoning.

8.3. Guilt

A further challenge arises from connecting the opportunity and disposal hypotheses—which we have split into identity- and event-level sub-hypotheses—to the guilt hypothesis. It is helpful to distinguish two inferential routes here. First, recall that the video, ID and SUV evidence, together, place the defendant near the crime scene at the relevant time. At the event level, the video evidence supports the further claim that whoever was near the crime scene at that time was very likely the perpetrator. The timing and proximate location are doing much of the work here. The video evidence seems to bear not just on the proposition that someone was near the crime scene, but on the strength of the inference from being near the crime scene to committing the crime. Stated more formally, the first inferential route toward the defendant's guilt is as follows:

$$\frac{\boldsymbol{N(x) \Longrightarrow G(x) \qquad N(d)}}{\boldsymbol{G(d)}}$$

In other words, whoever was near the crime scene at the relevant time was likely the perpetrator, and the defendant was near the crime scene at the relevant time; therefore, the defendant was likely the perpetrator[11].

[11] The arrow⟹ is meant to signal the defeasible relation between antecedent ***N(x)*** and consequent ***G(x)***. Another approach would use a a conditional probability statement, such as ***P(G(x) | N(x))***, and assign it a reasonably high value.

The second inference toward guilt goes through the disposal route. The phone location evidence links the defendant to the location where the firearm was recovered. The ballistics evidence establishes that the recovered firearm was the firearm used in the shooting. That fact supports a bridge from disposal to guilt: whoever later disposed of that firearm was likely the shooter. Stated more formally,

$$\frac{D(x,l) \Rightarrow G(x) \qquad D(d,l)}{G(d)}$$

The two routes converge on the same conclusion. One moves from the defendant's presence near the crime scene to guilt; the other from the defendant's presence at the location where the firearm was disposed of to guilt.

As before, we could hard code these inferences in a Bayesian network by adding suitable conditional probability tables. But that move would require assigning a defeasible conditional statement to a node in the network. In principle, this is possible: a node can stand for any proposition, however complex. But the cost is that the structure of the inference can no longer be read off the network.

A related oddity concerns the video evidence and the ballistics evidence, which was not formally modeled. These items of evidence appear to support a relation between propositions, rather than simply the truth of a single proposition. The video evidence supports the inference from someone's being near the crime scene at the relevant time to that person's being the shooter. The ballistics evidence supports the inference from someone's disposing of the firearm to that same person's having committed the shooting. In standard propositional Bayesian networks, however, nodes usually represent atomic propositions or events, while inferential relations among propositions are represented through the graph structure and the conditional probability tables. So evidence that bears on the strength of an inferential connection does not have an obvious place in the model[12].

9. CONCLUSION

If the analysis in the previous section is on the right track, the expressive limitation I have identified is not with probability theory as such. The problem lies instead with the expressive resources of standard, propositional Bayesian networks. Their nodes normally stand for atomic propositions, but some legally important inferences have a richer logical structure. The inference about opportunity, for example, depends on identity: if $\boldsymbol{v} = \boldsymbol{d}$, then predicates true of $\boldsymbol{v}$ may be transferred to $\boldsymbol{d}$. Other

[12] It is in principle possible to add further nodes so that the inferential relation is reified as another proposition. The model can then become formally workable, but such a move would obscure the fact that the evidence is about an inference from one proposition to another.

inferences depend on conditional claims, such as the claim that whoever was near the crime scene at the relevant time was the shooter. As we have seen, a standard, propositional Bayesian network can approximate such inferences by hard coding the relevant relationships into conditional probability tables. But that is only a workaround. Although it is workable in simple cases, one wonders whether this approach breaks down as the legal case to be modeled grows more complex.

A better option would be to rely on a formal framework with greater expressive power. If probability could be combined with a first-order language—made of variables, predicates, conditionals, and identity—then many of the problematic inferences we have examined could be represented more directly. Instead of treating ***v = d***, ***N(v)***, and ***N(d)*** as separate propositional nodes connected by an *ad hoc* table, the model could represent the general identity principle that, if two terms co-refer, predicates true of one are true of the other. Many evidential inferences in legal cases concern people, objects, locations, and events whose identities are uncertain. A satisfactory model should be able to represent inferences as to whether the person in a video is the defendant; whether the firearm recovered at one location is the firearm used in the shooting; whether the person who disposed of the firearm is the same person who appeared near the crime scene; and so on.

The challenge, then, is how probability should be combined with a first-order language. To be sure, there is already a body of work in this direction. Multi Entity Bayesian Networks (MEBNs), for example, were developed as a first-order language for probabilistic reasoning (Laskey, 2008). But MEBNs were not developed for legal applications, so the feasibility of using them to model legal evidence deserves further examination. That is a task for another time.

REFERENCES

Allen, R. J., & Pardo, M. S. (2007). The Problematic Value of Mathematical Models of Evidence. *Journal of Legal Studies*, *36*(1), 107–140.

Bozza, S., Taroni, F., & Aitken, C. (2026). The Myth of Judicial Objectivity: Why Subjectivity is the only Path to a Transparent Trial. *Quaestio Facti. Revista Internacional Sobre Razonamiento Probatorio*, *11*.

Dahlman, C., & Kolflaath, E. (2022). Causal Models versus Reason Models in Bayesian Networks for Legal Evidence. *Synthese*, *200*, 477.

Fenton, N., Neil, M., & Lagnado, D. A. (2013). A General Structure for Legal Arguments about Evidence using Bayesian Networks. *Cognitive Science*, *37*(1), 61–102.

Fenton, N., Neil, M., Yet, B., & Lagnado, D. A. (2020). Analyzing the Simonshaven case using Bayesian Networks. *Topics in Cognitive Science*, *12*(4), 1092–1114.

Hampson, L., & Leeuwen, L. van. (2025). Investigating the Value of Qualitative Bayesian Networks of Complete Cases as "Double-Check" Tools on Traditional Judicial Reasoning: An Exploratory Study. *Proceedings of the Workshop on AI for Evidential Reasoning*, *4157*, 40–71. https://ceur-ws.org/Vol-4157/paper5.pdf

Kadane, J. B., & Schum, D. A. (1996). *A Probabilistic Analysis of the Sacco and Vanzetti Evidence*. John Wiley & Sons.

Lagnado, D. A. (2021). *Explaining the Evidence: How the Mind Investigates the World*. Cambridge University Press.

Laskey, K. B. (2008). MEBN: A Language for First-Order Bayesian Knowledge Bases. *Artificial Intelligence*, *172*(2–3), 140–178.

Mackor, A. R. (2026). Bayesian Modelling of Criminal Cases as a Whole: A Philosophical Reflection on Dutch Case Law. *Quaestio Facti. Revista Internacional Sobre Razonamiento Probatorio*, *10*, 363–384.

Stacey, J., Opperman, S., Curran, J. M., & Harbison, S. A. (2025). Current State and Barriers to Global Adoption of Forensic Evaluative Reporting of Findings Given Activity-Level Propositions. *Science & Justice*, *65*(6), 101336.

Taroni, F., Biedermann, A., Bozza, S., Garbolino, P., & Aitken, C. (2014). *Bayesian Networks for Probabilistic Inference and Decision Analysis in Forensic Science* (2nd ed.). John Wiley & Sons.

Taylor, D., Biedermann, A., Hicks, T., & Champod, C. (2018). A Template for Constructing Bayesian Networks in Forensic Biology Cases when Considering Activity Level Propositions. *Forensic Science International: Genetics*, *33*, 136–146.

Tribe, L. H. (1971). Trial by mathematics: Precision and Ritual in the Legal Process. *Harvard Law Review*, *84*(6), 1329–1393. https://doi.org/10.2307/1339610

Vlek, C. S., Prakken, H., Renooij, S., & Verheij, B. (2014). Building Bayesian Networks for Legal Evidence with Narratives: A Case Study Evaluation. *Artificial Intelligence and Law*, *22*(4), 375–421.

Quaestio facti. Revista Internacional sobre Razonamiento Probatorio
Quaestio facti. International Journal on Evidential Legal Reasoning
Sección: Conjeturas y refutaciones
2026 I 11 pp. 181-198
Madrid, 2026
DOI: 10.33115/udg_bib/qf.i11.23280
Marcial Pons Ediciones Jurídicas y Sociales
© Mario Günther
ISSN: 2604-6202
Recibido: 07/05/26 | Aceptado: 16/06/26 | Publicado *online*: 26/06/2026
Editado bajo licencia Reconocimiento 4.0 Internacional de Creative Commons

BAYESIAN DECISION THEORY CAN GUIDE LEGAL FACTFINDING

Mario Günther*
Panthéon Sorbonne

ABSTRACT: I argue that Bayesian decision theory can guide legal factfinding. I do so by offering an account of legal proof on which judges should minimize expected justice costs. My account entails a judge's credence threshold for finding guilty and his prior credence of guilt. Hence, it can guide a judge in his decision based on the lawful evidence presented at trial—unlike the Bayesian model Mackor presents.

KEYWORDS: Philosophy of Law, Legal Proof, Bayesianism, Decision Theory, Retributive Justice, Undeserved Punishment.

SUMMARY: 1. INTRODUCTION.— 2. THE PRINCIPLE OF EXPECTED UTILITY MAXIMIZATION.— 3. JUSTICE COSTS: 3.1. A Simpler Credence Threshold; 3.2. The Severity of Undeserved Punishment; 3.3. The Severity of the Crime and Appropriate Punishments.— 4. BAYESIAN EVIDENCE EVALUATION: 4.1. How to Determine the Prior Credence of Guilt?; 4.2. The Strength of the Evidence Required for Conviction.— 5. CONCLUSION.— APPENDIX.— REFERENCES

* Please contact Mario.Gunther@univ-paris1.fr for any inquiries. ORCID: https://orcid.org/0000-0001-6208-448X.

1. INTRODUCTION

Mackor (2026) asks whether 'the' Bayesian model can and should guide the judicial evaluation of evidence in criminal cases as a whole. She focuses on Dutch criminal law, which allows a judge to convict a defendant for an offense only if the judge is convinced that the defendant committed it in light of the lawful evidence presented at trial. However, a judge may not convict even if the presented lawful evidence convinces him that the defendant is guilty. Mackor points out that this freedom of choice leaves the judge "helpless" in his decision (p. 364). For he lacks further guidance on whether or not he should convict the defendant. Dutch criminal law leaves open when the presented lawful evidence is sufficient for a finding of guilt. It leaves open how the judge should decide.

The Bayesian decision criterion is often conceptualized as a threshold for the posterior probability that the defendant committed the alleged offence—the final probability that the defendant is guilty after learning all the lawful evidence presented at trial (Günther & Friedrich, 2026). However, a judge has only a posterior of guilt if he starts out with a prior probability of guilt—the probability of guilt with which a judge begins to evaluate the lawful evidence received at trial. Mackor suggests that the judge's prior of guilt could be based on some reference class. This poses the question what reference class of the many possible ones should be used to estimate the prior (fn.4). If there is no suitable reference class, she suggests relegating the task of determining the prior of guilt to "the court"—without giving any further guidance (pp. 366-7&fn. 11).

Surprisingly, Mackor does not offer a Bayesian standard of proof in terms of a probability threshold. She neither says when a final probability of guilt suffices for finding a defendant guilty, nor what a judge's prior of guilt should be in general. The Bayesian model Mackor presents leaves open how a judge should decide. As it stands, the model cannot guide judges in their decisions.

Here I argue that Bayesian decision theory can guide a judge's factfinding. I do so by offering a decision-theoretic account of legal proof which does not leave open how a judge should decide. The account is based on the heart of Bayesian decision theory—the principle of expected utility maximization (EUM). But it goes beyond the formalism of decision theory by making a substantive assumption about the justice costs of the possible trial outcomes. The normative assumption entails with the principle of EUM a probability threshold for finding guilty from which I derive a prior probability of guilt. The result is a Bayesian account for how a judge should decide in criminal trials. The account shows that Bayesian decision theory can guide legal factfinding in principle.

The plan is as follows. Section 2 explains how the principle of EUM can figure as the decision criterion in legal factfinding. Section 3 argues for an assignment of utilities or justice costs to the possible trial outcomes based on the substantive as-

sumption that a judge's decision-making only aims at establishing retributive justice and avoiding undeserved punishment. The result is a simplified probability threshold for finding guilty in terms of the severity of undeserved punishment. Section 4 outlines what probability assessments a judge should have based on the total lawful evidence presented at trial, how a judge's prior of guilt should be determined, and how strong the total lawful evidence must be for a conviction. This last step allows us to complete the model Mackor presents such that it coincides with ours, and so does not leave open how a judge should decide.

2. THE PRINCIPLE OF EXPECTED UTILITY MAXIMIZATION

How should a judge decide? Bayesian decision theory answers that any agent should maximize her expected utility (Jeffrey, 1983; Ramsey, 1926; Savage, 1972; von Neumann & Morgenstern, 1944). This principle of expected utility maximization (EUM) says that the choices of an agent should solely depend on her value and probability assessments. The value assessments of an agent are her preferences. The probability assessments of an agent are her degrees of belief, or credences for short. If an agent's preferences satisfy certain rationality axioms, her preferences can be represented by a utility function and her credences by a probability function. A rational agent's expected utility of a choice option is the sum of her utilities of the possible outcomes weighted by her credences that the respective outcomes are true. The principle of EUM says that a rational agent chooses an option that has the highest expected utility among the available options.

Let us consider the decision situation a rational judge faces in a criminal trial as a whole. As Mackor (2026, p. 368) says, the judge needs to decide between the ultimate hypothesis *G* that the defendant is guilty and the ultimate hypothesis *I* that the defendant is not guilty or 'innocent' based on the lawful evidence presented at trial. The two ultimate hypotheses are negations of each other, and so they mutually exclude each other and jointly exhaust the space of logical possibilities. The judge has the options of finding guilty (FG) or not (FI). Hence, the possible outcomes of the trial are a true finding of guilt (*TG*), a false finding of guilt (*FG*), a true finding of not guilty (*TI*), and a false finding of not guilty (*FI*). The judge's decision can be summed up in the following decision table:

	guilty (*G*)	not guilty (*I*)
finding guilty (FG)	*TG*	*FG*
finding not guilty (FI)	*FI*	*TI*

The principle of expected utility maximization provides a decision criterion: the judge should find guilty if and only if (iff) doing so has higher expected utility than finding not guilty:

$$EU(\mathrm{FG}) > EU(\mathrm{FI}).^{1} \qquad (1)$$

The expected utility of finding guilty is:

$$EU(\mathrm{FG}) = P(G) \cdot U(TG) + P(I) \cdot U(FG).$$

The expected utility of not finding guilty is:

$$EU(\mathrm{FI}) = P(G) \cdot U(FI) + P(I) \cdot U(TI).$$

A strength of decision theory is that it allows to determine credence thresholds for finding guilty in terms of the utilities. Proposition 1 of the Appendix shows that the inequality (1) is equivalent to:

$$P(G) > \frac{U(TI) - U(FG)}{U(TG) - U(FG) + U(TI) - U(FI)} = \theta. \qquad (2)$$

The principle of expected utility maximization recommends a judge to find guilty just in case his credence $P(G)$ of guilt meets the threshold θ obtained from the utility values he assigns to the possible outcomes. I make a proposal for what utility values a judge should assign in the next section.

3. JUSTICE COSTS

A judge should make a just decision. Bayesian decision theory on its own does not say what decisions are just. But it offers a framework for spelling out one's notion of justice in terms of a judge's utility assignments to the possible trial outcomes (Kaplan, 1968; Kaye, 1980, 1982, 1999). I will do so by making the substantive assumption that a judge's decision has only two aims: to establish retributive justice and to avoid undeserved punishment. My assumption implies that *all* of the consequences unrelated to the two aims that a judge's decision may have should not matter for his utility assignments. It is substantive by excluding consequences, such as the social benefit of deterring and incapacitating dangerous or recidivism apt criminals. Indeed, these consequences are unrelated to establishing retributive justice for the case under consideration and avoiding undeserved punishment for the defendant. My substantive assumption leads to specific costs to justice, as I will explain in what follows.

Retributive justice makes two core claims. First, wrongdoers should be punished only in virtue of having done wrong. Second, wrongdoers deserve punishment in proportion to the harm they caused—in proportion to the severity of the offense they committed. So the severity of the punishment should be proportional to the

[1] Strictly speaking, the principle of expected utility maximization allows a judge both to find either guilty or not guilty if both findings have the same expected utility. We stipulate that in such rare cases of ties, a judge should find not guilty.

severity of the wrong done. In sum, the fundamental idea of retributive justice is this: a person deserves punishment iff she committed a wrongful act; and she deserves a more severe punishment for more severe wrongful acts. To punish an innocent person and not to punish a guilty person are therefore deviations from establishing retributive justice.

In the context of a criminal trial, the defendant deserves punishment iff she is guilty of committing a crime; and she deserves punishment in proportion to the severity of the crime. Unless otherwise stated, I assume that the punishment is proportional to the severity of the crime and so appropriate. On this assumption, a judge's true findings establish retributive justice, whereas his false findings do not. Any defendant deserves a true finding of not guilty and the appropriate punishment of a true finding of guilt. By contrast, no defendant deserves a false finding of not guilty and the punishment incurred by a false finding of guilt. The latter trial outcome incurs undeserved punishment on the defendant and so violates the judge's second aim to avoid such.

By my substantive assumption, there is only one way for a judge's decision to be unjust: the decision does not establish retributive justice. And only an unjust finding of guilt incurs undeserved punishment to the defendant. Let the *justice cost* of a possible trial outcome measure the extent to which it is unjust. We can think of costs to justice as deviations from establishing retributive justice weighted by the severity of undeserved punishment. In particular, I propose a two-tier account of how unjust a possible trial outcome is:

(i) A possible trial outcome bears a cost of justice iff it is unjust.

(ii) The justice cost of an unjust trial outcome is greater in proportion to the severity of undeserved punishment it incurs on the defendant.

(i) says that a possible trial outcome bears a cost of justice iff it does not establish retributive justice for the case under consideration. Defendants deserve the trial outcomes of true findings but not those of false findings. A false finding of not guilty incurs a justice cost, but it does not incur an undeserved punishment to the defendant for the simple reason that there is no punishment in findings of not guilty. A false finding of guilt incurs a justice cost and an undeserved punishment for the defendant. By (ii), false findings of guilt bear a greater cost to justice than false findings of not guilty because the severity of some undeserved punishment is always higher than that of no punishment.

I have made a rough proposal for how to measure the justice costs of possible trial outcomes. This allows us to think of a judge's decision as follows: he should minimize the expected justice costs. I will refine and integrate the proposal into my decision-theoretic account in the next two subsections.

3.1. A Simpler Credence Threshold

What justice costs should a judge assign to true findings? On the assumption that the severity of the punishment is proportional to that of the crime, a judge's true findings establish retributive justice perfectly: a true finding of not guilty establishes retributive justice perfectly and exactly as much as a true finding of guilt. Hence, true findings bear no cost to retributive justice and incur no undeserved punishment.

The justice costs for each possible trial outcome are the respective utility assignments. There are no justice costs of true findings:

$$U(TG) = U(TI) = 0.$$ [2]

As a result, the formula to determine the credence threshold for finding guilty in equation (2) simplifies to:

$$P(G) > \frac{-U(FG)}{-U(FG) - U(FI)} = \theta.$$ [3] (3)

On my account, the credence threshold is always above 1/2. For this to be seen, note that a false finding of guilt incurs some punishment for the defendant and a false finding of not guilty does not. Some undeserved punishment is always more severe than none. So the justice cost of a false finding of guilt is greater than that of a false finding of not guilty. Hence, my account excludes the case where the credence threshold for finding guilty is below 1/2. This case means trouble for other decision-theoretic accounts: it says that a judge should find guilty if his credence meets the low credence threshold—even if the judge thinks it less likely that the defendant is guilty than not (Günther, 2024b, p. 139). On my account, a rational judge never finds guilty when he thinks that the defendant is less likely to be guilty than not.

3.2. The Severity of Undeserved Punishment

What justice costs should the judge assign to false findings? To recap, any false finding does not establish retributive justice and so bears a cost of justice by (i) of my proposal. A false finding of not guilty does not incur any punishment on the defendant, and so no undeserved punishment. Its justice cost is by (ii) not greater than

[2] We can set the justice costs of true findings to zero without loss of generality because utility functions are only unique up to positive affine transformation:

$$U(O) = aU'(O) + b, \text{ where } a, b, \in \mathrm{R} \text{ and } a > 0$$

[3] Kaplan (1968) also considers only the justice costs or 'disutilities' of false findings but without the argument from retributive justice. He has been criticized for doing so, among others, by Tribe (1971), Lillquist (2002), Laudan and Saunders (2009), and Nance (2016). However, their counterarguments assume that the utilities of true convictions may differ from those of true acquittals—contrary to our notion of retributive justice.

the baseline justice cost for false findings required by (i). As utility functions are only unique up to positive affine transformations, we can set the baseline cost to justice to

$$U(FI) = -1.$$

By doing so and setting the cost of true findings to zero, we only determine that a rational judge prefers any true finding over a false finding of not guilty.

A false finding of guilt incurs the baseline justice cost by (i) and is greater than that cost in proportion to the severity of undeserved punishment it incurs on the defendant by (ii). Let us represent the severity of undeserved punishment by a severity factor s. The severity factor s equals one if there is no undeserved punishment and is strictly greater than one if there is undeserved punishment. As the baseline cost to justice is represented by the cost of a false finding of not guilty, we obtain:

$$U(FG) = U(FI) \cdot s, \text{ where } s \geq 1. \qquad (4)$$

The severity of undeserved punishment is nothing but the ratio of the justice cost of a false finding of guilt to that of a false finding of not guilty.

The second tier of my proposal is motivated by a principle of minimizing undeserved harm: a judge should minimize the undeserved punishment to the defendant. Hence, the justice cost of a false finding of guilt grows with the amount of undeserved punishment it incurs on the defendant. On my account, the severity of an undeserved punishment can explain why falsely convicting a defendant of murder bears a greater justice cost than falsely convicting a defendant of shoplifting expensive headphones. Murder is a more severe crime than shoplifting, and so murderers deserve a harsher punishment than shoplifters by retributive justice.

As a matter of fact, law codes determine the severity of punishments against defendants for different crimes. Dutch law, for example, codifies that murder is punishable by life imprisonment or a fixed-term prison sentence of up to 30 years, while shoplifting is only punishable up to 4 years in prison and a fine up to €27.500. An imprisonment for 30 years is a more severe punishment than an imprisonment for 4 years plus the fine. By (ii), the justice cost to falsely imprison a defendant for 30 years is greater than to falsely imprison a defendant for 4 years plus the fine. A false finding in a murder case therefore has a greater justice cost than a false finding in a shoplifting case—also in the Netherlands.

Note the asymmetry to false findings of not guilty. There is no punishment in false acquittals, and so its justice cost cannot scale with the severity of undeserved punishment to the defendant. My account has this general upshot: a judge's justice cost of falsely convicting a defendant increases with the severity of the undeserved punishment borne by the defendant; however, his cost of a false acquittal does not.

On my account, the credence threshold for a finding of guilt is raised by a higher justice cost of false findings of guilt, but it is never lowered by the justice cost of false

findings of not guilty because the latter cost remains constant for a given crime[4]. It follows that my account allows us to express the credence threshold in terms of the severity factor alone. By substituting in (3) the justice cost $U(FG)$ of a false finding of guilt with that $U(FI)$ of a false finding of not guilty weighted by the severity s of the undeserved punishment, we obtain:

$$P(G) > \frac{-U(FI) \cdot s}{-U(FI) \cdot s - U(FI)} = \frac{-s}{-s-1} = \frac{s}{s+1}. \qquad (5)$$

The more severe the undeserved punishment is, the higher the credence threshold for finding guilty. It is only the severity of the possible undeserved punishment borne by the defendant that shifts the credence threshold for finding guilty.

3.3. The Severity of the Crime and Appropriate Punishments

My account offers the means to derive the magnitude of the justice costs from the appropriate severity of punishments for crimes. By retributive justice, the severity of punishment should depend on how severe the crime is. Stealing a bottle of water is a less severe crime than stealing a large amount of money or committing criminal tax fraud. In many law codes, the former is punishable by a small fine, while the latter is punishable by jail. If we assume that these law codes are appropriate, the severity of the punishment is directly proportional to the severity of the crime. If so, the severity of the punishment, or equivalently of the crime, can be used to determine the appropriate justice costs, and so to determine the appropriate credence threshold for finding guilty.

On my account, the credence threshold for finding guilty varies with the appropriate severity of the punishment. If Dutch law, for example, encodes the appropriate severity of a punishment for a crime, we can derive how much greater the justice cost of a false finding of guilt for the crime is compared to that of a false finding of not guilty. For illustration, suppose we appropriately measure the severity of punishment in terms of years in jail as follows: any 3 years in jail add a unit of severity, and one year in prison corresponds to a fine of €500.000. A 30-years imprisonment for murder gives us then a severity factor of 10. This result is reminiscent of the famous thought by Blackstone (1753, p. 358) that "the law holds that it is better that ten guilty persons escape than that one innocent suffer." If this is an appropriate severity assessment, we obtain that the justice cost $U(FG)$ of a false finding of guilt is ten times greater than the baseline justice cost $U(FI)$. And so the credence threshold for a murder finding is 10/11, or approximately 0.91.

[4] Ross (2023, pp. 1090-1, 1097) comes to a similar conclusion even though he rejects "popular" decision-theoretic accounts of the standard of proof.

A 4-years imprisonment plus a fine of €27.500 gives us a severity factor of $4/3 + (1/3 \cdot 27.500/500.000) \approx 1.352$. If this severity assessment is appropriate, we obtain that the justice cost $U(FG)$ of a false finding of guilt is roughly 1.352 times greater than the baseline cost $U(FI)$. And so the credence threshold for a theft finding is roughly 0.57. This is, of course, only a proof of concept. It is tricky to find a good unit of severity and translations between years in jail and monetary fines. The main issues are how the severity of the punishment should be measured and what the exact functional form of severity growth should be—linear, logarithmic, quadratic, exponential, some combination, or something else altogether. I leave an appropriate assessment of the unit of severity and its functional form of growth for future research.

We have seen how appropriate amounts of punishment can be translated into justice costs of false findings of guilt on my account. My account thereby offers a method to estimate the appropriate ratio of the justice costs for a given crime from its severity. This being said, my account can also be used in the reverse direction. If we know the ratio of justice costs for a given crime, we can determine its appropriate severity. Hence, my account is compatible with a method of reflective equilibrium between a crime's justice costs, its severity, and the credence threshold for finding guilty (Jellema, 2025).

4. BAYESIAN EVIDENCE EVALUATION

We have seen that taking the principle of expected utility maximization as a decision criterion for a judge implies that his standard of proof is equivalent to a credence threshold. Given the appropriate justice costs, a judge should find the defendant guilty iff his credence in the defendant's guilt meets the threshold in light of the lawful evidence presented at trial.

The open question is what credences a judge should have based on the lawful evidence presented at trial? Decision theory only stipulates that a rational agent's credences conform to the probability axioms. Bayesianism adds that she updates her credences by conditionalization or a generalization thereof. There are no further epistemic constraints on a rational agent imposed by Bayesian decision theory.

A rational judge determines his credence of guilt by updating on the total lawful evidence presented at trial. A rational judge starts out with a prior credence function P_0, which represents his first credences before learning any evidence presented at trial[5]. He conditionalizes his prior credence $P_0(G)$ that the defendant is guilty on the total lawful evidence E_T to obtain the final credence $P(G)$ of guilt:

$$\boldsymbol{P(G) = P_0(G \mid E_T).} \qquad \textbf{(6)}$$

[5] Strictly speaking, the prior credence function $P_0(\cdot) = P_{00}(\cdot \mid B)$ encodes some background information B about the world but no evidence that is relevant to the hypothesis G that the defendant is guilty in the absence of the evidence presented at trial.

The conditional probability is given by the ratio definition of conditional probability:

$$\boldsymbol{P_0(G \mid E_T) = \frac{P_0(G \cap E_T)}{P_0(E_T)}}.^{6} \qquad (7)$$

The total lawful evidence E_T is equivalent to the conjunction of all individual pieces of lawful evidence $E_1 \cap E_2 \cap \ldots \cap E_n$ successively presented at trial. It is easy to see that a rational judge's final credence of guilt is the same whether he conditionalizes on the total lawful evidence or on all individual pieces successively.

4.1. How to Determine the Prior Credence of Guilt?

A rational judge's final credence of guilt depends on his prior. For this to be seen, observe that $P_0(G \cap E_T) = P_0(E_T \mid G) \cdot P_0(G)$, and so by (7)

$$\boldsymbol{P_0(G \mid E_T) = \frac{P_0(E_T \mid G) \cdot P_0(G)}{P_0(E_T)}}. \qquad (8)$$

This formula is known as Bayes's Theorem. It is often used to calculate the posterior credence after learning some evidence because the likelihood $P_0(E_T \mid G)$ of the evidence given that the defendant is guilty is often easier to assess in practice than the credence $P_0(G \cap E_T)$ of the conjunction of guilt and evidence. Bayes's Theorem makes explicit that a rational judge's posterior and final credence of guilt depends on his prior.

For illustration, suppose the rational judge's likelihood of the total lawful evidence given that the defendant is guilty is 0.7, his prior credence of this evidence is 0.1, and his credence threshold for finding guilty is 0.9. Then his final credence of guilt is 0.91 for a prior of 0.13 and 0.84 for a prior of 0.12. The small difference in the prior credence of guilt determines whether or not his final credence meets the threshold. A rational judge's credence of guilt is sensitive to his prior. As long as there is not a unique prior, it often remains unclear whether or not a judge should find guilty—on the Bayesian model Mackor presents and any other. But legal decisions should not be arbitrary in this way (Dahlman, 2018, p. 19). This is the *problem of the prior* in the legal context.

---

[6] It follows that

$$P(E_T) = P_0(E_T \mid E_T) = \frac{P_0(E_T \cap E_T)}{P_0(E_T)} = 1.$$

For cases, where the rational judge does not become certain of the evidence after learning it, many Bayesians use Jeffrey conditionalization—a generalization of conditionalization due to Jeffrey (1983).

Opponents of Bayesian accounts of legal factfinding justifiedly take the problem of the prior as an argument against them. The prior credence of guilt cannot be determined by the lawful evidence presented at trial. The prior must already be in place before any evidence presented at trial is received by a rational judge at trial. Otherwise the Bayesian machinery does not get off the ground. Hence, the problem of the prior is perhaps the most serious problem for Bayesian accounts of legal factfinding.

What prior credence of guilt should a judge have? I think a fundamental principle of criminal procedure can help. The presumption of innocence says that a defendant should be considered innocent until proven guilty by the lawful evidence presented at trial. It implies that a defendant is to be considered not guilty until the evidence proves otherwise. The presumption of innocence thereby constrains the appropriate prior of guilt. On a credence threshold view, it says at least this: a rational judge presumes a defendant 'innnocent' or not guilty at the beginning of the trial iff his prior credence $P_0(G)$ of guilt is below the credence threshold θ for finding guilty. If $P_0(G) < \theta$ and no evidence were presented, the rational judge would consider the defendant innocent in the sense of finding not guilty. The problem with this notion of the presumption of innocence is that it does not single out a prior credence of guilt. Indeed, it leaves way too many candidates for the prior credence of guilt on the table—infinitely many.

I propose to understand the presumption of innocence as follows: a rational judge's prior credence of guilt should be $P_0(G) = 1 - \theta$, where θ is the credence threshold for finding guilty. It follows that the higher the credence threshold for a finding of guilt, the lower a rational judge's prior credence of guilt, and so the more demanding it is to find the defendant guilty. As θ is always strictly above 1/2 on my account, a rational judge's prior of guilt is always strictly below 1/2. Indeed, a rational judge should at least assume that the defendant is less likely to be guilty than not at the outset of a trial. This ensures that a rational judge never finds a defendant guilty without considering at least some lawful evidence.

My method of setting the prior credence of guilt is again motivated by the severity of (the appropriate punishment for) the crime under dispute. The more severe the crime, the higher the credence threshold for finding guilty and the lower the prior credence of guilt. My account entails that a judge should presume a defendant not guilty to the degree that the crime is severe. My notion of the presumption of innocence figures as a safeguard for not incurring undeserved punishment on the defendant, just like the credence threshold for finding guilty does. And so my Bayesian account for the evaluation of evidence got off the ground.

My method for setting a normative prior credence of guilt determines a *presumed prior* in the sense of Stein (1996), Posner (1999), Dahlman (2018), and Dahlman and Kolflaath (2021). My presumed prior follows Friedman's (2000) criticism in rejecting Posner's reading of the presumption of innocence as a prior credence of guilt of 1/2. I also part way with Dahlman and Kolflaath who assume without argument that the prior credence of guilt must be much smaller than 1/2 (p. 296). My

presumed prior is determined by how high the credence threshold for a finding of guilt should be in terms of justice costs. Hence, it is not arbitrary.[7] My standard of proof is determined by both, the prior credence of guilt and the credence threshold for finding guilty.

4.2. The Strength of the Evidence Required for Conviction

My account sets a rational judge's credence threshold θ for a finding of guilt and his prior credence $P_0(G) = 1 - \theta$ of guilt. So we can think of the standard of proof in terms of how strong the total lawful evidence presented at trial must be to increase the prior credence of guilt such that it meets the credence threshold. As Dahlman (2018, p. 26) shows, the two numeric values allow us to compute the strength of the evidence required for a conviction if the strength is measured by the likelihood ratio of the total lawful evidence. This ratio compares the likelihood of the total lawful evidence if the defendant is guilty to that if the defendant is not guilty:

$$\frac{P_0(E_T \mid G)}{P_0(E_T \mid \neg G)}. \qquad (9)$$

In the Appendix, I prove a theorem that refines Dahlman's result in the confines of my account:

Theorem 1 A rational judge's final credence of guilt meets the threshold $P_0(G \mid E_T) > \theta$ for a finding of guilt iff

$$\frac{P_0(E_T \mid G)}{P_0(E_T \mid \neg G)} > \frac{\theta^2}{(1-\theta)^2}.$$

On my account of legal proof, a rational judge finds a defendant guilty iff his likelihood ratio of the total lawful evidence is greater than his ratio of the squared credence

[7] I largely follow Dahlman and Kolflaath's (2021) arguments against determining priors of guilt on empirical grounds, such as the number of possible perpetrators (Fenton *et al.*, 2017; Lindley, 1977). The main problem is that empirical facts make the priors somewhat arbitrary: one judge might assign an empirical prior of guilt based on one reference class for a case, where another would assign a higher or lower such prior based on a different reference class. Reichenbach's (1949) reference class problem is looming large here (Colyvan *et al.*, 2001; Hájek, 2007). Building on Reichenbach, Franklin (2011, pp. 559-61) proposes that the ideal reference class for finding guilty or not should be defined by all features correlated with the defendant's guilt as long as the probability estimate of guilt remains reliable. This seems to counter the charge that reference classes are arbitrary in principle. But they are still arbitrary in practice because what features are correlated with the guilt hypothesis depends on the available data. More importantly, there are hypothetical cases, where even an ideal reference class might give us a prior of guilt which surpasses the threshold for finding guilty. No evidence is then needed to find a defendant guilty. And so the empirical prior based on the ideal reference class conflicts with the presumption of innocence.

threshold for finding guilty to the squared prior credence of guilt. For illustration, suppose that the credence threshold for finding guilty is 0.9. Then a rational judge finds the defendant guilty iff he thinks the total lawful evidence presented at trial is 81 times more likely if the defendant is guilty than if the defendant is not guilty.

We are now in a position to complete the Bayesian model Mackor (2026) presents. The model is the odds formulation of Bayes's Theorem of equation (8):

$$\frac{P_0(G \mid E_T)}{P_0(I \mid E_T)} = \frac{P_0(G)}{P_0(I)} \cdot \frac{P_0(E_T \mid G)}{P_0(E_T \mid I)}. \qquad (10)$$

The model allows us to calculate the posterior odds in terms of the likelihood ratio of the total lawful evidence and the prior odds. If the likelihood ratio is greater (lower) than one, the posterior odds are greater (lower) than the prior odds. The likelihood ratio measures the upward or downward impact of the total lawful evidence on the posterior odds. As Mackor and Prakken (2026) know, the model allows us to deduce the final probability of guilt but only because the two ultimate hypotheses exclude each other and jointly exhaust the logical space—as Proposition 2 in the Appendix shows.

One can use my account to establish a normative credence threshold for a finding of guilt, and so a normative prior credence of guilt. If one uses the odds formulation of Bayes's Theorem together with the threshold θ and the prior $1 - \theta$ of guilt, the package coincides with my account of when a judge should find a defendant guilty. As $P_0(G \mid E_T) > \theta$ iff $P_0(I \mid E_T) < 1 - \theta$, we obtain $P_0(G \mid E_T) > \theta$ iff

$$\frac{P_0(G \mid E_T)}{P_0(I \mid E_T)} > \frac{1 - \theta}{\theta} \cdot \frac{\theta^2}{(1 - \theta)^2} = \frac{\theta}{1 - \theta}.$$

Hence, my account has the resources to turn the model Mackor presents into an account of legal proof, which does not leave open how a judge should decide.

5. CONCLUSION

I offered a Bayesian account of legal proof on which a judge's decision should minimize expected justice costs. The principle of expected utility maximization and my notion of justice costs entail a credence threshold for finding guilty in terms of the severity of undeserved punishment. The prior credence of guilt is derived from the normative credence threshold and implements a presumption of innocence. The normative prior and credence threshold determine how strong the lawful evidence presented at trial must be for a finding of guilt. They also make the model Mackor presents equivalent to mine, and so capable of guiding a judge through a whole criminal trial.

I have left several issues for future work. Among them are to explicitly show that my account covers civil cases as well; to defend the normative foundations of Bayesian decision theory for legal factfinding (Fenton & Lagnado, 2021); to defend my specific notion of justice costs against others; to examine the extent to which actual judges can be guided in their decisions by the regulative ideal of a rational judge; and to tackle open problems for decision-theoretic accounts, such as the proof paradoxes (Günther, 2024a, 2024c; Redmayne, 2008). But my account does show that Bayesian decision theory can guide a rational judge in his decision and so can serve as a guide to legal factfinding in principle.

APPENDIX

Proposition 1

The inequality $EU(\mathrm{FG}) > EU(\mathrm{FI})$ is equivalent to (2) in the main text.

Proof. Suppose $EU(\mathrm{FG}) > EU(\mathrm{FI})$. By the definition of expected utility, we obtain

$$P(G) \cdot U(TG) + P(I) \cdot U(FG) > P(G) \cdot U(FI) + P(I) \cdot U(TI),$$

which is equivalent to

$$P(G) \cdot U(TG) - P(G) \cdot U(FI) > P(I) \cdot U(TI) - P(I) \cdot U(FG).$$

By the Law of Total Probability, $P(I) = 1 - P(G)$. So

$$P(G) \cdot U(TG) - P(G) \cdot U(FI) > (1 - P(G)) \cdot U(TI) - (1 - P(G)) \cdot U(FG).$$

By multiplying out the right-hand side, we obtain

$$P(G) \cdot U(TG) - P(G) \cdot U(FI) > U(TI) - P(G) \cdot U(TI) - U(FG) + P(G) \cdot U(FG)$$

Rearranging all probability weighted utilities to the left-hand side gives us

$$P(G) \cdot U(TG) - P(G) \cdot U(FI) + P(G) \cdot U(TI) - P(G) \cdot U(FG) > U(TI) - U(FG),$$

and so

$$P(G) \cdot (U(TG) - U(FI) + U(TI) - U(FG)) > U(TI) - U(FG).$$

Dividing both sides by $U(TG) - U(FI) + U(TI) - U(FG)$ results in

$$P(G) > \frac{U(TI) - U(FG)}{U(TG) - U(FI) + U(TI) - U(FG)}$$

Proposition 2

The final probability of guilt can be calculated from the prior odds and the likelihood ratio of the total lawful evidence only because the two ultimate hypotheses of guilt G and not guilty $\neg G$ are mutually exclusive and jointly exhaustive.

Proof. By Bayes's Theorem

$$P_0(G \mid E_T) = \frac{P_0(G) \cdot P_0(E_T \mid G)}{P_0(E_T)}.$$

Only because G and $\neg G$ are mutually exclusive and jointly exhaustive, we obtain by the Law of Total Probability

$$P_0(E_T) = P_0(G) \cdot P_0(E_T \mid G) + (1 - P_0(G)) \cdot P_0(E_T \mid \neg G).$$

Substituting $P_0(E_T)$ in Bayes's Theorem yields

$$P_0(G \mid E_T) = \frac{P_0(G) \cdot P_0(E_T \mid G)}{P_0(G) \cdot P_0(E_T \mid G) + (1 - P_0(G)) \cdot P_0(E_T \mid \neg G)}.$$

Divide the numerator and denominator by $P_0(E_T \mid \neg G)$:

$$P_0(G \mid E_T) = \frac{P_0(G) \cdot \dfrac{P_0(E_T \mid G)}{P_0(E_T \mid \neg G)}}{P_0(G) \cdot \dfrac{P_0(ET \mid G)}{P_0(E_T \mid \neg G)} + (1 - P_0(G))}.$$

This shows that the final probability of guilt can be calculated from the prior probability of guilt and the likelihood ratio of the total lawful evidence if the two ultimate hypotheses are G and $\neg G$. However, observe that the derivation only works because the two ultimate hypotheses are mutually exclusive and jointly exhaustive. Otherwise, one cannot substitute $P_0(E_T)$ in Bayes's Theorem using the Law of Total Probability. Similarly, one cannot derive the prior probability of guilt from the odds ratio alone unless $P_0(\neg G) = 1 - P_0(G)$.

In general, some posterior odds

$$\frac{P_0(H_1 \mid E_T)}{P_0(H_2 \mid E_T)}.$$

do not give us the posterior probability $P_0(H_1 \mid E_T)$. To obtain the posterior probability, one would need to know the value x of the posterior odds and that $P_0(H_2 \mid E_T) = 1 - P_0(H_1 \mid E_T)$. Only if so:

$$P_0(H_1 \mid E_T) = x \cdot (1 - P_0(H_1 \mid E_T)).$$
$$P_0(H_1 \mid E_T) = x - x \cdot P_0(H_1 \mid E_T).$$
$$P_0(H_1 \mid E_T) + x \cdot P_0(H_1 \mid E_T) = x.$$
$$P_0(H_1 \mid E_T) \cdot (1 + x) = x.$$
$$P_0(H_1 \mid E_T) = \frac{x}{1+x}.$$

Theorem 1. A rational judge's final credence of guilt meets the threshold $P_0(G \mid E_T) > \theta$ for a finding of guilt iff

$$\frac{P_0(E_T \mid G)}{P_0(E_T \mid \neg G)} > \frac{\theta^2}{(1 - \theta)^2}.$$

Proof. Suppose $P_0(G \mid E_T) > \theta$. By Bayes's Theorem and the Law of Total Probability, we obtain

$$\frac{P_0(G) \cdot P_0(E_T \mid G)}{P_0(G) \cdot P_0(E_T \mid G) + (1 - P_0(G)) \cdot P_0(E_T \mid \neg G)} > \theta.$$

By multiplying both sides with the denominator, we get

$$P_0(G) \cdot P_0(E_T \mid G) > \theta \cdot P_0(G) \cdot P_0(E_T \mid G) + \theta \cdot (1 - P_0(G)) \cdot P_0(E_T \mid \neg G).$$

By subtracting the term $\theta \cdot P_0(G) \cdot P_0(E_T \mid G)$ on both sides, we obtain

$$P_0(G) \cdot P_0(E_T \mid G) - \theta \cdot P_0(G) \cdot P_0(E_T \mid G) > \theta \cdot (1 - P_0(G)) \cdot P_0(E_T \mid \neg G),$$

which we can rewrite as

$$(1 - \theta) \cdot P_0(G) \cdot P_0(E_T \mid G) > \theta \cdot (1 - P_0(G)) \cdot P_0(E_T \mid \neg G).$$

Dividing both sides by $(1 - \theta) \cdot P_0(G)$ and $P_0(E_T \mid \neg G)$ gives us

$$\frac{P_0(E_T \mid G)}{P_0(E_T \mid \neg G)} > \frac{\theta \cdot (1 - P_0(G))}{(1 - \theta) \cdot P_0(G)}.$$

On my account, $P_0(G)$ is defined as $1 - \theta$. Note that we can go through the entire proof backwards. Hence, we obtain the desired result.

REFERENCES

Blackstone, W. (1753). *Commentaries on the laws of england: In four books* (Vol. 2). J.B. Lippincott.

Colyvan, M., Regan, H. M., & Ferson, S. (2001). Is it a crime to belong to a reference class. *Journal of Political Philosophy*, *9*(2), 168–181. *https://doi.org/10.1111/1467-9760.00123*

Dahlman, C. (2018). Determining the base rate for guilt. *Law, Probability and Risk*, *17*(1), 15–28. *https://doi.org/10.1093/lpr/mgx009*

Dahlman, C., & Kolflaath, E. (2021). The problem of the prior in criminal trials. In *Philosophical foundations of evidence law* (pp. 287–300). Oxford: Oxford University Press.

Fenton, N., & Lagnado, D. (2021, September). Bayesianism: Objections and rebuttals. In *Philosophical foundations of evidence law*. Oxford University Press. *https://doi.org/10.1093/oso/9780198859307.003.0019*

Fenton, N., Lagnado, D., Dahlman, C., & Neil, M. (2017). The opportunity prior: A simple and practical solution to the prior probability problem for legal cases. *Proceedings of the 16th Edition of the International Conference on Articial Intelligence and Law*, 69–76. *https://doi.org/*10.1145/3086512.3086519

Franklin, J. (2011). The objective Bayesian conceptualisation of proof and reference class problems. *Sydney Law Review*, *33*(3), 545–561.

Friedman, R. D. (2000). A presumption of innocence, not of even odds. *Stanford Law Review*, *52*(4), 873–887.

Günther, M. (2024a). Epistemic sensitivity and evidence. *Inquiry*, *67*(6), 1348–66. *https://doi.org/10.1080/0020174X.2021.1936158*

Günther, M. (2024b). Legal proof should be justified belief of guilt. *Legal Theory*, *30*(3), 129–141. *https://doi.org/10.1017/S1352325224000089*

Günther, M. (2024c). Probability of guilt. *Canadian Journal of Philosophy*, *54*(3), 189–206. *https://doi.org/10.1017/can.2025.6*

Günther, M., & Friedrich, C. (2026). A Bayesian justification for the scenario approach to legal proof. *Proceedings of the Workshop on AI for Evidential Reasoning (AI4EVIR 2025)*, *4157*.

Hájek, A. (2007). The reference class problem is your problem too. *Synthese*,
156(3), 563–585. *https://doi.org/10.1007/s11229-006-9138-5*

Jeffrey, R. C. (1983). *The logic of decision* (Second). University of Chicago Press.

Jellema, H. (2025). Standards of proof for future crimes and decision theory. *Quaestio facti. International Journal on Evidential Legal Reasoning*, (10). *https://doi.org/10.33115/udg* bib/qf.i10. 23144

Kaplan, J. (1968). Decision theory and the factfinding process. *Stanford Law Review*, *20*(6), 1065–1092.

Kaye, D. H. (1980). Naked statistical evidence. *The Yale Law Journal*, *89*(3), 601–611.

Kaye, D. H. (1982). The limits of the preponderance of the evidence standard: Justifiably naked statistical evidence and multiple causation. *American Bar Foundation Research Journal*, *7*(2), 487–516.

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), 1–28. *https://doi.org/10.1177/136571279900300101*

Laudan, L., & Saunders, H. D. (2009). Rethinking the criminal standard of proof: Seeking consensus about the utilities of trial outcomes. *International Commentary on Evidence*, *7* (2). *https://doi.org/doi:10.2202/1554-4567.1099*

Lillquist, E. (2002). Recasting reasonable doubt: Decision theory and the virtues of variability. *UC Davis Law Review*, *36*, 85–197.

Lindley, D. V. (1977). Probability and the law. *Journal of the Royal Statistical Society. Series D (The Statistician)*, *26*(3), 203–220.

Mackor, A. R. (2026). Bayesian modelling of criminal cases as a whole: A philosophical reflection on Dutch case law. *Quaestio facti. International Journal on Evidential Legal Reasoning*, (10). *https://doi.org/10.33115/udg bib/qf.i10.23229*

Mackor, A. R., & Prakken, H. (2026). On reporting likelihood ratio's of exhaustive and non-exhaustive hypotheses about rare events in criminal cases. *Proceedings of the Workshop on AI for Evidential Reasoning (AI4EVIR 2025)*, *4157*.

Nance, D. A. (2016). *The burdens of proof: Discriminatory power, weight of evidence, and tenacity of belief*. Cambridge University Press.

Posner, R. A. (1999). An economic approach to the law of evidence. *Stanford Law Review*, *51*, 1477–1546.

Ramsey, F. P. (1926). Truth and probability. In R. Braithwaite (Ed.), *The foundations of mathematics and other logical essays* (1999 electronic edition, pp. 156–198). Harcourt.

Redmayne, M. (2008). Exploring the proof paradoxes. *Legal Theory*, *14*(4), 281–309.

Reichenbach, H. (1949). *The theory of probability*. University of California Press.

Ross, L. (2023). Criminal proof: Fixed or flexible? *The Philosophical Quarterly*, *73*(4), 1077–1099. *https://doi.org/10.1093/pq/pqad001*

Savage, L. J. (1972). *The foundations of statistics* (Second Revised). Wiley Publications in Statistics.

Stein, A. (1996). Judicial factfinding and the bayesian method: The case for deeper scepticism about their combination. *The International Journal of Evidence & Proof*, *1*(1), 25–47. *https://doi.org/10.1177/136571279600100103*

Tribe, L. H. (1971). Trial by mathematics: Precision and ritual in the legal process. *Harvard Law Review*, *84*(6), 1329–1393.

von Neumann, J., & Morgenstern, O. (1944). *Theory of games and economic behavior*. Princeton University Press.

Quaestio facti. Revista Internacional sobre Razonamiento Probatorio
Quaestio facti. International Journal on Evidential Legal Reasoning
Sección: Conjeturas y refutaciones
2026 I 11 pp. 199-223
Madrid, 2026
DOI: 10.33115/udg_bib/qf.i11.23303
Marcial Pons Ediciones Jurídicas y Sociales
© Leya Lisa Hampson
ISSN: 2604-6202
Recibido: 08/06/2026 | Aceptado: 15/06/2026 | Publicado *online*: 30/06/2026
Editado bajo licencia Reconocimiento 4.0 Internacional de Creative Commons

COMBATTING BAYESIAN CRITICISM: A BAYESIAN-INSPIRED CRITICAL CHECKLIST FOR JUDICIAL REASONING

Leya Lisa Hampson [1]

ABSTRACT: This paper develops a Bayesian-inspired checklist of critical questions designed to support probabilistic reasoning in legal contexts. Building on Mackor's (2026) proposal, rather than requiring formal Bayesian models or numerical probabilities, the framework translates key Bayesian principles into a practical sequence of guided questions aimed at helping judges improve and structure their reasoning and avoid probabilistic fallacies. is intended to assist judges in structuring and critically reflecting on their reasoning, while helping to identify and avoid probabilistic fallacies.

The framework addresses common reasoning errors discussed in the legal literature (see Dahlman, 2023), including base-rate neglect, inversion fallacies, false dichotomies, dependence neglect, convergence neglect, and link-skipping. Organised according to key stages of probabilistic reasoning, the checklist is designed as a flexible aid to judicial deliberation and self-evaluation rather than a formal decision-making model. It aims to promote transparency in reasoning, encourage explicit reflection on assumptions and evidential relationships, and support the critical evaluation of expert evidence. Future research will focus on empirical testing and further refinement of the framework.

KEYWORDS: Bayesian reasoning, Legal aid, Probabilistic fallacies

SUMMARY:

1. INTRODUCTION: 1.1. The Problem: Probabilistic Reasoning in Judicial Contexts; 1.2. Suggested solutions: The Bayesian Method.— 2. METHOD: FROM

[1] Department of Transboundary Legal Studies, Faculty of Law, University of Groningen

THEORY TO STRUCTURED CHECKLIST.— 3. CHECKLIST DEVELOPMENT: 3.1. Framing the Hypotheses and Evidence; 3.2. Prior Probabilities and Base Rate Neglect: 3.2.1. Motive and Opportunity; 3.3. False dichotomies and competing hypotheses; 3.4. Evidence Evaluation: Inversion Fallacies; 3.5. Evidence Evaluation: Convergent Evidence, Dependencies and their Combination; 3.6. Indirect Reasoning and Chains of Inference; 3.7. Instructions for application in practice.— 4. CONCLUSION.— REFERENCES.— APPENDIX A: PROPOSED CHECKLIST .— APPENDIX B: PROPOSED INFORMATION SHEET

1. INTRODUCTION

The central question addressed by Mackor (2026) is whether the Bayesian model can provide meaningful guidance for the judicial evaluation of evidence, in particular through the modelling of complete criminal cases. While this approach offers a formally coherent framework for reasoning under uncertainty, Mackor highlights the practical and conceptual concerns associated with its application in judicial contexts. These include discussions on the feasibility of constructing full Bayesian networks in real cases, as well as the ability of judges to apply such models reliably and independently.

In light of these limitations, Mackor identifies a number of directions for future research. Among these is the proposal to translate the insights of Bayesian reasoning into a structured checklist of "critical questions" that legal practitioners could use when analysing evidence to help them recognise and avoid probabilistic fallacies.

The present paper takes up this proposal directly. Rather than focusing on the construction of complete Bayesian models, it seeks to operationalise the suggested list of critical questions by developing a structured checklist for use in legal, in particular judicial, reasoning. In doing so, it builds on the insights of Bayesian analysis, while deliberately avoiding the need for numerical quantification or the construction of full Bayesian networks. The aim is to provide judges with a practical tool for identifying and avoiding probabilistic fallacies in their own reasoning, as well as for critically assessing reasoning presented in expert reports. In this way, the paper offers a concrete response to the research agenda outlined by Mackor, by developing and structuring the proposed checklist into an instrument that can be implemented in judicial (and broader legal practice). In its current stage, the research remains a theoretical, offering an initial version of the checklist in the literature. The framework will undergo further refinement through feedback and iterative development before being subjected to rigorous empirical evaluation in future studies.

1.1. The Problem: Probabilistic Reasoning in Judicial Contexts

Judicial decision-making involves reasoning under uncertainty. Judges are required to evaluate competing hypotheses in light of incomplete and often complex

evidence, including probabilistic information derived from expert testimony. A substantial body of research has demonstrated that, in such contexts, legal factfinders are susceptible to systematic errors in reasoning (Aitken & Taroni, 2004; Cole, 2004; Dahlman, 2015; Dahlman *et al.*, 2016; De Keijser, J. & Elffers, 2012; Fenton & Neil, 2000; Koehler, 1993; Koehler, 2008; Koehler & Thompson, 2006; Martire *et al.*, 2013; Pilditch *et al.*, 2019; Saks & Koehler, 2008; Thompson & Schumann, 1987). These include a range of probabilistic fallacies, including but not limited to, base-rate neglect, inversion errors, and the misinterpretation of convergent evidence, as identified in the literature (see Dahlman (2023) for a full overview).

Such errors are not merely theoretical concerns; they have been associated with documented miscarriages of justice such as in the case of Lucia de Berk (Gill *et al.*, 2010). They often arise not from a lack of information, but from failures to properly structure and evaluate the relationship between hypotheses and evidence. Importantly, such probabilistic fallacies result in flawed judicial reasoning even if the reasoning does not result in a formal miscarriage of justice (Mackor, 2022). From a normative perspective, the laws of probability provide a framework for coherent reasoning under uncertainty, helping to avoid internal contradictions and inconsistencies in fact-finding (Hunt and Mostyn, 2019).

The approach adopted in this paper therefore assumes that the uncertainty inherent in legal reasoning can, at least in principle, be expressed in probabilistic terms. At the same time, it recognises that the use of probability, more specifically the Bayesian model, in legal contexts remains contested. Some legal scholars and practitioners question whether legal reasoning can or should be quantified (Meester, 2020; Meester & Stevens, 2025). This tension highlights the need for approaches that retain the benefits of probabilistic reasoning while remaining accessible and commonly accepted within legal practice.

1.2. Suggested solutions: The Bayesian Method

Within the literature on rational legal proof in the Netherlands[2], three broad approaches are commonly discussed: Bayesian models, scenario-based (or narrative) approaches, and argumentation-based frameworks (Prakken, 2018). These differ, among other things, in their treatment of probabilistic reasoning. Bayesian and scenario-based approaches typically involve indirect probabilistic reasoning, in which hypotheses are evaluated in light of how well they explain the evidence, whereas argumentation-based approaches often involve more direct reasoning from evidence to conclusions.

The Bayesian method, in particular, offers a coherent framework for reasoning under uncertainty. A key advantage of the Bayesian approach is that it makes the role

[2] This is not limited to the Netherlands, see e.g., the Simonshaven case analysis (Prakken *et al.*, 2020).

of prior assumptions and the impact of new evidence on these assumptions explicit (Lagnado *et al.*, 2012). However, its application in legal contexts has been widely debated. A common criticism concerns the assignment of numerical probabilities, summarised in the question "where do the numbers come from?" (Tillers & Green, 1988; Meester & Stevens, 2024), as well as the complexity and required expertise involved in constructing and interpreting complete Bayesian networks (Prakken, 2020).

Recent research (Vlek *et al.*, 2013), although still limited, further suggests that the construction of complete Bayesian networks for entire criminal cases may not be feasible for judges in practice, due to both time constraints and the level of technical expertise required[3]. This issue was investigated by Hampson & Leeuwen (2026), who independently constructed Bayesian models based on the same criminal case verdict. Their findings indicate that the process of constructing such models yields important qualitative benefits, including greater transparency in reasoning, clearer identification of assumptions, and improved understanding of evidential relationships. Notably, the subsequent comparison and discussion of these models proved particularly insightful, as it helped to clarify the structure and implications of Bayesian reasoning.

Against this background, Mackor (2026) argues that certain objections to Bayesian analysis lose force when judges themselves engage with the method, for example during deliberations. In such cases, Bayesian reasoning need not function as a primary decision-making framework, but rather as a reflective tool. Judges may first arrive at a preliminary assessment of the evidence using more traditional methods and subsequently compare this assessment with the outcome of a more structured Bayesian analysis. In this sense, the method operates as a form of a post hoc inspection, enabling judges to critically evaluate their initial conclusions.

Taken together, this raises the question of whether similar benefits might be achieved through more accessible means, by providing structured support that guides reasoning without requiring the construction of full Bayesian models.

The following section builds on this by outlining how Bayesian principles can be implemented in a structured checklist of critical questions to support correct reasoning and safeguard against probabilistic fallacies in legal reasoning, In doing so, it seeks to retain the normative strengths of Bayesian reasoning, such as the explicit consideration of hypotheses, evidence, and assumptions, while avoiding common criticisms associated with formal Bayesian approaches, including the need for precise numerical probabilities or fully specified Bayesian networks. The checklist is designed as a tool for judicial self-evaluation, intended for both individual and collective use. Its purpose is to support judges in identifying and avoiding probabilistic fallacies in their own reasoning, as well as in critically assessing the reasoning presented in

[3] Although this requires further research, as discussed by Mackor (2026).

expert reports. By structuring reflection through guided questioning, the proposed framework aims to make key elements of probabilistic reasoning more transparent and accessible in practice.

2. METHOD: FROM THEORY TO STRUCTURED CHECKLIST

Building on the conceptual framework outlined above, this section sketches out the procedure through which the proposed checklist was developed. This is not presented as a formalised methodology, but rather as a theory-informed account of the considerations that shaped the design. The aim was to translate theoretical insights from Bayesian reasoning and literature on probabilistic fallacies into a structured and practically applicable instrument.

Thus, the underlying structure of the checklist is informed by Bayesian reasoning, although it is not applied in a formal or quantitative manner. Instead, it reflects key elements of Bayesian inference, such as the consideration of prior assumptions, the evaluation of the likelihood of evidence under competing hypotheses, and the comparison of alternative explanations, and follows the order in which these typically arise in Bayesian modelling[4].

Following Mackor (2026), when implementing a Bayesian framework, it is first necessary to clarify the conditions under which it is to be applied: In particular, three questions must be addressed: who performs the analysis, at what stage is it used, and what is the intended function. In the present case, these questions are answered as follows. First, the analysis is intended to be performed by legal professionals (judges, lawyers) themselves. Second, it is intended to guide deliberations in the court chamber as well as the write up of the judicial ruling. Third, its function is not to replace judicial reasoning, but to serve as a tool for reflection, enabling judges to critically assess, and where necessary, refine their evidential conclusions.

Alongside this, it is necessary to identify where judicial reasoning is prone to error. The checklist draws on a set of probabilistic fallacies identified in the legal literature, most notably the taxonomy proposed by Dahlman (2023). Dahlman list contains twelve fallacies organised into seven categories, of which 6 are discussed in this article: The product fallacy is not discussed here. This aligns with the aim of the checklist to avoid the focus on the numbers and remain usable for factfinders[5].

Each fallacy was treated as a point of departure for question design. The reasoning processes underlying these fallacies were examined, and, for each, a sequence of

[4] It should be noted that there is no universally accepted way to model Bayesian networks. This is based on both the legal idioms outlined by Lagnado *et al.* (2012) and the author's personal experience.

[5] The product fallacy refers to the failure to apply the product rule of probability theory to independent evidence supporting the same hypothesis, leading to an under-estimation of the combined probative value of the evidence (Dahlman, 2023). It inherently involves reasoning with numbers, therefore it is not included in the checklist.

questions was developed. These prompts were designed with the aim to counteract fallacious reasoning by uncovering the assumptions on which they rest, examining the evidential basis for those assumptions, and exploring alternative explanations or competing hypotheses.

Rather than consisting of isolated prompts, the checklist is structured as short sequences of interconnected questions, in which each question builds on the previous one. In this way, it aims to approximate a dialogue-like exchange (akin to discussions in the deliberation chamber amongst three judges) guiding the user step-by-step through the reasoning process whilst encouraging explicit reflection.

The question derivation process was further informed by reference to documented European cases in which such fallacies have occurred[6]. These cases were used as examples to illustrate how factfinders can arrive at erroneous conclusions in practice and to support the development of question sequences by targeting points where reasoning may break down.

Finally, given its intended application in real-world judicial contexts, focus was placed on (1) understandability, (2) transparency, and (3) feasibility. (1) The questions are formulated in clear, non-technical language and (2) are accompanied, where necessary, by brief explanations of relevant probabilistic concepts and fallacies. This is intended to ensure that the tool supports informed reasoning without imposing the technical demands associated with formal Bayesian modelling. (3) The checklist is designed to remain concise and practical within the time constraints of judicial reasoning. Questions are kept limited and organised into subgroups, facilitating the skipping of questions factfinders have already considered or regard as sufficiently addressed.

3. CHECKLIST DEVELOPMENT

Reasoning about legal evidence can be understood, at a basic level, as involving the evaluation of a hypothesis (H), typically whether the defendant is guilty or not, in light of observed evidence (E). While formal Bayesian networks provide a structured representation of these relationships, the present approach translates this structure into a sequence of guided questions. These questions are designed to guide the decision-maker through the key stages of probabilistic reasoning while identifying common sources of error. Accordingly, the checklist is organised into a series of Bayesian reasoning stages, drawing references to relevant probabilistic fallacies that may arise during at each stage.

[6] The Sally Clark and Lucia de Berk cases, amongst others, informed this derivation. Although these cases have already been discussed at length in the literature, they remain useful illustrations of how probabilistic reasoning errors may arise and be identified in practice. NB: The aim was not to provide an exhaustive review of all legal cases involving probabilistic fallacies, particularly as there may also be cases in which such fallacies were committed but not recognised or discussed. Rather, these cases are used as illustrative support for the development of the checklist and its underlying reasoning structure.

3.1, Framing the Hypotheses and Evidence

At the most basic level, probabilistic reasoning requires clarity as to what is being evaluated. Errors can arise where the hypothesis or the evidential basis is left implicit or insufficiently specified (Dahlman, 2023). Accordingly, the checklist begins by requiring the decision-maker to explicitly articulate both:

(1) *What is the main hypothesis under consideration?*

(2) *Have any alternative hypotheses been proposed? What are they?*

(3) *What evidence is available in the case?*

These initial questions establish the structure within which subsequent reasoning takes place, this is akin to the starting point of the construction of a Bayesian network (Lagnado *et al.*, 2012).

3.2. Prior Probabilities and Base Rate Neglect

A key feature of Bayesian inference is the consideration of the prior (Lagnado *et al.*, 2012). The prior refers to the probability of a hypothesis being true before any case-specific evidence is taken into account. Assuming the hypothesis concerns the defendant guilt, the prior probability is the probability of the defendant being guilty before considering any evidence. A simplified example may help illustrate this concept further. Suppose an inmate in a prison is accused of murdering another inmate. Due to the restricted location of the crime, investigators determine that only a limited group of individuals could have realistically committed the offence, including all prison staff and inmates with access to the area. If, for illustrative purposes, this group consists of 100 individuals, the initial probability that any one individual committed the crime is 1 in 100. In a formal Bayesian analysis, this represents the starting probability. Once additional evidence becomes available, this probability may increase or decrease.

Failure to consider this starting probability constitutes a well-documented probabilistic fallacy commonly referred to as 'base-rate neglect' (Bar-Hillel, 1980; Dahlman, 2023; Fenton and Neil, 2011; Gigerenzer, 1991; Koehler, 1996; Thomson and Schuman, 1987). The three following cases clearly illustrate this error.

The first example can be found in a Swedish case from 2019[7], in which an influencer was prosecuted for the murder of his girlfriend, who was found dead in a bathtub. Given that suicide is statistically far more common than homicide in Sweden, the municipal court's failure to adequately consider this base rate may have systematically biased reasoning in favour of the prosecution (Dahlman, 2023). The defendant

[7] Solna tingsrätt, 2019-04-04, B 3252-18

was initially found guilty before the verdict was later overturned by the Court of Appeals. The second example concerns the infamous case of Sally Clark, who was accused of murdering her two infant children. Central to the case was the comparison of two competing hypotheses: either the two babies died of sudden infant death syndrome (SIDS), or both children were murdered by their mother. During the trial, expert testimony focused heavily on the rarity of the two SIDS deaths occurring within the same family, presenting this event as extremely improbable. The fact that the probability of a mother killing both of her infant children is also extremely low, if not lower, was neglected. The final example refers to the case of Lucia de Berk, a Dutch nurse convicted for multiple murders. It was argued that it was a low probability that her proximity to multiple patient deaths was coincidental. Within this line of reasoning, the base rate of fatalities at the hospital, as well as nurses working on shift during such events was neglected. In all three cases, they were first convicted before the verdict was later overturned.

Although the concept of the prior is widely recognised as a central element of Bayesian reasoning, its operationalisation remains highly debated (Cheng & Pardo, 2015; Dahlman & Koflaath, 2021; Dawid, 2002; Friedman, 2002; Meester & Stevens, 2024; Posner, 1999; Stein, 1996; Sullivan, 2019). Various numerical approaches to determining priors have been proposed and explored in the literature (e.g., the opportunity prior (Fenton *et al.*, 2019)). At the same time, scholars have questioned whether such prior knowledge can always, or should always, be expressed in precise numerical form, particularly in legal contexts where the relevant background information may be limited (Meester & Stevens, 2024).

Given that much of the debate surrounding priors concerns the assignment of exact numerical values, the present framework deliberately avoids requiring judges to determine precise numerical probabilities. This is informed by both the literature and the three documented cases discussed above. In each of these cases, the central reasoning error was not necessarily that an incorrect numerical prior had been assigned, but rather that insufficient consideration had been given to the relative probability of competing explanations before the evidence was assessed. The core problem therefore concerned the weighing of competing hypotheses and the tendency to overlook relevant background probabilities altogether.

Accordingly, the proposed checklist does not focus on numerical quantification or provide guidelines on how to arrive at such, but rather on encouraging explicit reflection on whether competing hypotheses in their prior probability. Instead of asking judges to calculate exact probabilities, the checklist instead prompts consideration of which proposed explanation initially appears more probable and why.

(4) *Before considering any of the evidence, which hypothesis appears more probable?*

(4.1.) *Why do you consider this scenario more or less probable?*

3.2.1. Motive and Opportunity

The introduction of contextual factors, such as motive and opportunity, complicates this reasoning. Since motive and opportunity frequently shape the initial probability of competing explanations, these considerations are introduced at this stage of the checklist, before turning to the evaluation of specific items of evidence. In this way, the structure of the checklist follows the broader logic of Bayesian reasoning by first considering factors that may influence the relative plausibility of competing hypotheses, and only thereafter assessing the evidential value of individual pieces of evidence.

Although these factors are not equivalent to the prior probability, they influence the initial probability of competing hypotheses before the main evidential material is evaluated. While the unconditional prior probability of guilt may be difficult to determine, the introduction of motive or opportunity shifts the relevant question towards the conditional probability[8] of guilt given the presence of that motive. Making this distinction explicit helps clarify the role that such factors play in the overall argument (Lagnado *et al.*, 2012).

(4.2.) *Is there a motive? If so, what impact does the presence of this factor have on the probability of the scenario(s)?*

(4.3.) *Is there opportunity? If so, what impact does the presence of this factor have on the probability of the scenario(s)*

3.3. False dichotomies and competing hypotheses

Errors may also arise where factfinders incorrectly assume that competing hypotheses are both exhaustive and mutually exclusive. Hypotheses are exhaustive if they cover all possible outcomes or scenarios, meaning at least one must be true. Hypotheses are exclusive if only one can be true (if one is true, the other must be false). As discussed in the literature on argumentation, this "false dichotomy" can lead to the mistaken belief that evidence supporting one hypothesis necessarily undermines all alternatives to an equal degree (Dahlman, 2019; Govier, 2001).

In the Sally Clark case, two hypotheses were proposed: a double-murder or a double-case of SIDS. These hypotheses are neither exclusive nor exhaustive: They are not exclusive as the babies could have had SIDS and been murdered, these two hypotheses do not cancel each other out, or one baby could have died of SIDS and the other of murder, thus rendering both hypotheses partially true. They are not exhaustive, as they neglected to consider other possibilities such as one SIDS death and one murder, other medical causes, etc.

[8] The conditional probability of guilt refers to the probability a person is guilty, given specific evidence. It is calculated using Bayes Theorem, updating the prior probability of guilt based on new information.

To address this, the checklist includes prompts such as:

(5) *Are the competing hypotheses exhaustive (i.e., do they cover all possible outcomes, meaning that at least one must be true)?*

(6) *Are they mutually exclusive (i.e., if one is true does the other have to be false)?*

(7) *Does the evidence increase belief in more than one hypothesis at the same time?*

(8) *Is there an alternative explanation for the evidence?*

These questions are intended to prevent probabilistic fallacies, specifically false dichotomy, but also to broaden the scope of reasoning and prevent premature closure (see (8)).

3.4. Evidence Evaluation: Inversion Fallacies

Base-rate neglect rarely occurs in isolation and is often closely caused by another common probabilistic reasoning error: the inversion fallacy. In this fallacy, the probability of the evidence given the hypothesis, $P(E|H)$, is mistakenly equated with the probability of the hypothesis given the evidence, $P(H|E)$. As noted in the literature, a factfinder may for example reason that because the observed evidence is consistent with guilt, the probability of guilt must therefore be high. However, these probabilities are not equivalent, and such reasoning fails to consider the probability of observing the same evidence if the hypothesis were false (Dahlman, 2023).

The crucial question is not simply whether the evidence is consistent with the hypothesis, but whether the hypothesis provides a better explanation for the evidence than competing alternatives. In other words, the task is to assess whether the evidence is more likely to occur if the hypothesis is true than if it is false. This assessment can only be made once the competing hypotheses and their prior plausibility have been considered. Thus, the evaluation of evidence does not occur in isolation, but directly builds upon the prior reasoning stages presented above.

To showcase this fallacy in practice, we return to the cases of Sally Clark and Lucia de Berk. In the former, the jury was presented with statistical evidence by an expert witness suggesting that the occurrence of two SIDS cases within the same family was extremely unlikely. On this basis, the presence of two such deaths was treated as strong evidence against the defendant. The rarity of the observed event was effectively taken to indicate the unlikelihood of the defendant's innocence, as though the improbability of two natural deaths were equivalent to the improbability that the deaths occurred without foul play. However, these probabilities are not the same. The relevant question was not how unlikely the evidence is in isolation, but whether the competing hypothesis of a double homicide provided a more probable explanation for the deaths than the hypothesis of natural causes. Similarly, in the Lucia de Berk case, the court falsely assumed that the low probability of the deaths occurring by chance was equivalent to the probability that Lucia was innocent.

To reduce the risk of such inversion errors, the checklist explicitly focuses on the explanatory relationship between hypotheses and evidence. Rather than asking whether the evidence appears incriminating itself, the questions are designed to encourage reflection on whether the hypothesis adequately explains the observed evidence and whether the evidence can also reasonably occur if the hypothesis were false.

(9) *For each piece of evidence you wish to take into account, consider:*

(9.1.) *(How much) would you expect to see the observed evidence if the hypothesis were true?*

(9.2.) *(How much) would you expect to see the observed evidence if the hypothesis were false?*

(9.3.) *Is your expectation of seeing the evidence stronger if H is true or if H is false?*

The purpose of these questions is to ensure that (1) both true and false positive rates are considered, thereby aiming to restore the correct evidential balance and (2) reasoning proceeds from H to E, rather than directly from the existence of the evidence to the truth of the hypothesis.

3.5. Evidence Evaluation: Convergent Evidence, Dependencies and their Combination

In many cases, the strength of the prosecution's case rests not on a single piece of evidence, but on the combined effect of multiple items of evidence. Factfinders need to evaluate multiple pieces of evidence supporting the same hypothesis, commonly referred to as convergent evidence. However, integrating such evidence correctly can be challenging and result in various fallacies (Dahlman, 2023).

Evidence is dependent when the probative value of one item of evidence is affected by the existence of another (e.g., two witnesses who have talked to each other), independent evidence retains its probative value regardless of the presence of other evidence (e.g., two witnesses who have not been in contact with each other). A failure to recognise the cumulative effect of independent evidence may lead to underestimation, referred to as 'convergence neglect', while ignoring dependencies may lead to overestimation, this is known as 'dependence neglect' (Dahlman, 2023). For example, where several independent pieces of evidence each have a low probability of occurring if the hypothesis is false, their combined effect may be significantly stronger than any individual piece alone. Conversely, treating dependent evidence as independent can artificially inflate the perceived strength of the case.

Dependence neglect was committed in the Sally Clark case, where the expert witness fallaciously treated the occurrence of two SIDS cases within the same family as independent. If one child suffered from SIDS, the risk of the second child having SIDS increases, therefore the probability of the second infant dying from SIDS is

dependent on whether the first child died from SIDS. Documented cases[9] in which the dependency neglect fallacy has been committed commonly reflect fallacious assumptions and reasoning patters within expert analyses: This highlights the need for an additional layer of scrutiny when expert evidence is presented.

Another relevant risk in the evaluation of convergent evidence is the "zero-sum fallacy" (Dahlman, 2023; Pilditch *et al.*, 2019). This fallacy occurs where factfinders incorrectly assume that weakening one side of the evidential balance necessarily strengthens the opposing side to an equal degree. In probabilistic reasoning, however, evidential relationships are often non-linear and asymmetrical. The reduction in probative value of one piece of incriminating evidence does not automatically imply a corresponding increase in the probability of innocence, particularly where other independent evidence remains unaffected. Similarly, exculpatory evidence does not necessarily "cancel out" incriminating evidence of differing strength or reliability. Such reasoning risks distorting the cumulative assessment of evidence and may lead to both over- and underestimation of the overall strength of the case.

To address these issues, the checklist incorporates prompts aimed both at assessing the combination of evidence and at identifying potential dependencies:

(10) *Are any pieces of evidence dependent on one another (i.e., is the probative value of one piece of evidence influenced by the existence of another)?*

(10.1.) *If so, what makes them dependent and how might this relationship affect their probative value?*

(10.2.) *Has statistical expert evidence been presented in the case? If so, are any of the factors or pieces of evidence relied on by the expert potentially dependent?*

(10.3.) *Have any pieces of evidence been left out due to irrelevance or deemed as too small in probative value? Why?*

(10.4.) *Is there any evidence that probative value may have been overestimated or underestimated? Why?*

These questions are particularly important in complex cases involving multiple strands of evidence.

3.6. Indirect Reasoning and Chains of Inference

Finally, the checklist addresses cases in which evidence supports a hypothesis only indirectly, through a chain of inferences[10]. Such reasoning introduces additional opportunities for error, particularly where intermediate assumptions are not made

[9] Here, "cases" refers to cases commonly discussed in the European literature such as Sally Clark and Lucia de Berk. It should be noted that the author did not conduct a thorough systematic or exhaustive review of all relevant cases.

[10] This is often referred to as indirect or circumstantial evidence (Dahlman, 2023).

explicit. This is known as 'link-skipping' (Dahlman, 2023). This fallacy is best explained at hand of an example: Consider a case in which the defendant is accused of burglary. A positive ID is made between blue hoodie fibres found at the crime scene and a blue hoodie owned by the defendant. Claiming that the positive ID provides direct support for the defendant's guilt would be link-skipping, as it neglects the possibility (the sub-hypothesis) that someone else was wearing the defendant's hoodie at the crime scene. Committing this fallacy discounts part of the evidence uncertainty chain, thus causing an underestimation of the probability for a false positive. This can in turn lead to an overestimation of the probative value of the evidence chain.

To address this, the checklist includes prompts such as:

(11) *Does this piece of evidence relate to the hypothesis directly or through one or more inferential steps?*

(11.1.) *What assumptions are required to link this evidence to the hypothesis?*

(11.2.) *How strong or well-supported are these assumptions?*

(11.3.) *Is there an alternative explanation for each step in this chain?*

3.7. Instructions for application in practice

The checklist is intended to function as a flexible reasoning aid rather than a rigid procedural framework. It consists of a series of critical questions organised into thematic subgroups (for example, question 11, followed by questions 11.1., 11.2. and 11.3.). Practitioners should work through the checklist sequentially, reflecting on whether the relevant reasoning issue has been adequately considered within the case.

If a practitioner considers that a question has already been sufficiently discussed or reasoned through, that question may be skipped. Where a primary question has already been sufficiently addressed during reasoning or deliberation, practitioners may skip that question and its corresponding sub-questions. For example, if Question 11 is regarded as adequately considered, Questions 11.1.-11.3. may also be skipped.

The checklist is not intended to require extensive written responses or formal legal analysis under each question. Rather, the questions are designed to re-prompt critical reflections and encourage structured reasoning, drawing attention to potentially problematic probabilistic reasoning patterns that may otherwise go unnoticed. For the full checklist and detailed instructions for practitioners please see Appendix A.

4. CONCLUSION

This paper responds directly to the research agenda identified by Mackor (2026) by developing a structured checklist of critical questions aimed at improving probabilistic reasoning in judicial contexts. Rather than advocating for the direct use of

formal Bayesian models, the proposed framework translates key elements of Bayesian reasoning into a sequence of critical questions, thereby offering a practical and accessible tool for structured "self-evaluation".

The checklist is diagnostic in nature: it is designed to make reasoning processes transparent, to expose implicit assumptions, and to highlight potential sources of error. In this sense, it does not function as a decision-making tool and but as a mechanism for disciplined reflection. By enforcing a structured form of reflection in the reasoning process, it aims to reduce the probability that probabilistic fallacies remain unnoticed.

The scope of this approach is deliberately limited. It is not intended to resolve substantive uncertainties, nor to substitute for the expertise required to interpret complex statistical evidence. Such interpretive functions are already performed by forensic experts and advisors. The present framework aims to support judges in critically engaging with the reasoning presented and in recognising when additional expertise is required. Recognising that the probabilistic analysis of a piece of evidence is not understood, or that an assumption does not seem to have been justified, is itself a critical step towards more reliable decision-making.

At this stage, the proposed checklist still requires empirical validation. Future research will therefore focus on testing its effectiveness in practice, through experimental studies using legal scenarios in both third- and first-person formats (inspired by the work of Dahlman, 2025 and Hampson *et al.*, in progress). This dual format allows for the assessment of both the ability to identify reasoning errors and the ability to avoid such errors in one's own judgements.

This research forms part of the broader "Preventing Miscarriages of Justice" research project funded by the NWO (Grant number: 406.21.RB.004). The author thanks all affiliated members of the PMJ project for their extensive prior research and valuable input, which laid the foundations for the present work. Special thanks are also extended to the participants of the PMJ legal reasoning course in March and April 2026, whose engagement, questions and curiosity informally inspired many of the ideas developed in this paper, and the participants of the evidence workshop "Putting Evidence on the Stand" for their feedback.

REFERENCES

Aitken, C., & Taroni, F. (2004) Statistics and the Evaluation of Evidence, 2nd ed. Chichester: Wiley.

Bar-Hillel, M. (1980). The base-rate fallacy in probability judgments. *Acta Psychologica*, *44*(3), 211–233.

Cheng, E. K., & Pardo, M. S. (2015) 'Accuracy, Optimality and the Preponderance Standard', Law, Probability and Risk 193.

Cole, S. (2004) Grandfathering evidence: Fingerprint admissibility rulings from Jennings to Llera Plaza and back again. American Criminal Law Review 41: 1189–1276.

Dahlman, C. (2015). The felony fallacy. *Law, Probability and Risk*, *14*(3), 229–241. *https://doi.org/10.1093/lpr/mgv004*

Dahlman, C. (2019). De-Biasing Legal fact-finders with bayesian thinking. *Topics in Cognitive Science*, *12*(4), 1115–1131. *https://doi.org/10.1111/tops.12419*

Dahlman, C. (2023). A systematic account of probabilistic fallacies in legal fact-finding. *The International Journal of Evidence & Proof*, *28*(1), 45–64. *https://doi.org/10.1177/13657127231209019*

Dahlman, C. (2025). Training legal fact-finders to recognize probabilistic fallacies. *Law, Probability and Risk*, *24*(1). *https://doi.org/10.1093/lpr/mgaf010*

Dahlman, C., & Kolflaath, E.(2021) 'The Problem of the Prior in Criminal Trials', in C.

Dahlman, G. Tuzet and A. Stein (eds) Philosophical Foundations of Evidence Law. Oxford: OUP.

Dahlman, C., Zenker, F., & Sarwar, F. (2016). Miss rate neglect in legal evidence. *Law, Probability and Risk*, *15*(4), 239–250. *https://doi.org/10.1093/lpr/mgw007*

Dawid, A. P. (2002) 'Bayes' Theorem and Weighting Evidence by Juries', in Swinburne R. (ed.) Bayes Theorem, Proceedings of the British Academy, p. 113

De Keijser, J., & Elffers, H. (2012). Understanding of forensic expert reports by judges, defense lawyers and forensic professionals. Psychology, Crime & Law, 18 (2), 191-207

Fenton, N., & Neil, M. (2000) The jury observation fallacy and the use of Bayesian networks to present probabilistic legal arguments. Mathematics Today 37: 61–102.

Fenton, N., Lagnado, D., Dahlman, C., & Neil, M. (2019). The opportunity prior: A proof-based prior for criminal cases. *Law, Probability and Risk*. *https://doi.org/10.1093/lpr/mgz007*

Fenton, N., & Neil, M. (2011) Avoiding Probabilistic Reasoning Fallacies in Legal Practice using Bayesian Networks. AUJlLegPhil 5, 36 Australian Journal of Legal Philosophy. 114.

Fenton, N., Neil, M., & Lagnado, D. A. (2012). A general structure for legal arguments about evidence using Bayesian networks. *Cognitive Science*, *37*(1), 61–102. *https://doi.org/10.1111/cogs.12004*

Friedman, R. (2002) 'A POI—Not of Even Odds', Stanford Law Review 52

Gigerenzer, G. (1991). From tools to theories: A heuristic of discovery in cognitive psychology. *Psychological Review*, *98*(2), 254–267. *https://doi.org/10.1037/0033-295x.98.2.254*

Gill, R. D., Groeneboom, P., & de Jong, P. (2018). Elementary statistics on trial—the case of Lucia de Berk. *CHANCE*, *31*(4), 9–15. *https://doi.org/10.1080/09332480.2018.1549809*

Govier, T. (2001). *A practical study of argument*. Wadsworth/Cengage Learning.

Hampson, L., & Van Leeuwen, L. (2026) Investigating the value of qualitative Bayesian networks of complete cases as "double-check" tools on traditional judicial reasoning: An exploratory study. Proceedings of the AI4EVIR Workshop on AI for Evidential Reasoning, Turín, Italia.

Hunt, I., & Mostyn, J. (2019). Probability reasoning in judicial fact-finding. *The International Journal of Evidence & Proof*, *24*(1), 75–94. *https://doi.org/10.1177/1365712719875753*

Koehler, J. (1993) Error and exaggeration in the presentation of DNA evidence at trial. Jurimetrics 21: 21–40.

Koehler, J. J. (1996). The Base Rate Fallacy Reconsidered: Descriptive, Normative, and methodological challenges. *Behavioral and Brain Sciences*, *19*(1), 1–17. *https://doi.org/10.1017/s0140525x00041157*

Koehler, J. (2008) Fingerprint error rates and proficiency tests: What they are and why they matter. Hastings Law Journal 59: 1077–1100.

Koehler, J. J., & Thompson, W. C. (2006). Mock jurors' reactions to selective presentation of evidence from multiple-opportunity searches. *Law and Human Behavior*, *30*(4), 455–468. *https://doi.org/10.1007/s10979-006-9021-4*

Mackor, A. R. (2026). Bayesian modelling of criminal cases as a whole. *Quaestio Facti. Revista Internacional Sobre Razonamiento Probatorio*, (10). *https://doi.org/10.33115/udg_bib/qf.i10.23229*

Mackor, A. R., & Schutgens, R. (2022). De kansendeskundige als lekenrechter: Artikel 116 Grondwet. En B. Rijpkema & A. Ellian (Eds.), Een nieuw commentaar op de grondwet (pp. 451-466). Boom Publishers

Martire, K. A., Kemp, R. I., Watkins, I., Sayle, M. A., & Newell, B. R. (2013). The expression and interpretation of uncertain forensic science evidence: Verbal equivalence, evidence strength, and

the weak evidence effect. *Law and Human Behavior*, *37*(3), 197–207. *https://doi.org/10.1037/lhb0000027*

Meester, R. (2019). The limits of Bayesian thinking in Court. *Topics in Cognitive Science*, *12*(4), 1205–1212. *https://doi.org/10.1111/tops.12478*

Meester, R., & Stevens, L. (2024). Bayesian reasoning and the prior in court: Not legally normative but unavoidable. *Law, Probability and Risk*, *23*(1). *https://doi.org/10.1093/lpr/mgae001*

Pilditch, T. D., Fenton, N., & Lagnado, D. (2018). The zero-sum fallacy in evidence evaluation. *Psychological Science*, *30*(2), 250–260. *https://doi.org/10.1177/0956797618818484*

Posner, R. (1999) An Economic Approach to the Law of Evidence Stanford Law Review, 51.

Prakken, H. (2018a). A new use case for argumentation support tools: Supporting discussions of bayesian analyses of complex criminal cases. *Artificial Intelligence and Law*, *28*(1), 27–49. *https://doi.org/10.1007/s10506-018-9235-z*

Prakken, H. (2018b). A new use case for argumentation support tools: Supporting discussions of bayesian analyses of complex criminal cases. *Artificial Intelligence and Law*, *28*(1), 27–49. *https://doi.org/10.1007/s10506-018-9235-z*

Prakken, H. (2019). An argumentation-based analysis of the Simonshaven case. *Topics in Cognitive Science*, *12*(4), 1068–1091. *https://doi.org/10.1111/tops.12418*

Prakken, H., Bex, F., & Mackor, A. R. (2020). Editors' review and introduction: Models of rational proof in criminal law. *Topics in Cognitive Science*, *12*(4), 1053–1067. *https://doi.org/10.1111/tops.12519*

Saks, M., & Koehler, J. (2008) The Individualization Fallacy in Forensic Science Evidence. Vanderbilt Law Review 61: 199–220.

Solna tingsrätt, 2019-04-04, B 3252-18.

Stein, A. (1996) 'Judicial Fact-finding and the Bayesian Method', International Journal of Evidence & Proof, 1.

Sullivan, S. P, (2019) 'A Likelihood Story: The Theory of Legal Fact-Finding', University of Colorado Law Review, 1.

Thompson, W. C., & Schumann, E. L. (1987). Interpretation of statistical evidence in criminal trials: The prosecutor's fallacy and the defense attorney's fallacy. *Law and Human Behavior*, *11*(3), 167–187. *https://doi.org/10.1007/bf01044641*

Tillers, P., & Green, E. D. (1988). *Probability and inference in the law of evidence: The uses and limits of bayesianism*. Kluwer Academic Publishers.

Vlek, C., Prakken, H., Renooij, S., & Verheij, B. (2013) Modeling crime scenarios in a Bayesian network. In Proceedings of the Fourteenth International Conference on Artificial Intelligence and Law (ICAIL '13). Association for Computing Machinery, New York, NY, USA, 150–159.

APPENDIX A: PROPOSED CHECKLIST

Instructions for use during reasoning process:

Review each checklist item step-by-step, either individually or as a group. For each item, ask whether it has been considered adequately. Answer the questions verbally, cognitively, or through discussion; written responses are not required.

Instructions for use during deliberation:

Use the checklist as a reflective tool during discussions. Review each checklist item step-by-step as a group. Pause on any item that feels uncertain or under-con-

sidered and continue once it has been sufficiently addressed. Responses should occur through discussion rather than formal written completion.

Additional guidance:

If you become stuck on a question or are uncertain about the meaning of a term or concept, please refer to the accompanying information sheet for clarification. Examples have also been provided to aid the application.

Checklist

1. What is the main hypothesis under consideration?

2. Have any alternative hypotheses been proposed? What are they?

3. What evidence is available in the case?

4. Before considering any of the evidence, which hypothesis appears more probable?

4.1. Why do you consider this scenario more or less probable?

4.2. Is there a motive? If so, what impact does the presence of this factor have on the probability of the scenario(s)?

4.3. Is there opportunity? If so, what impact does the presence of this factor have on the probability of the scenario(s)

5. Are the competing hypotheses exhaustive (i.e., do they cover all possible outcomes, meaning that at least one must be true)?

6. Are they mutually exclusive (i.e., if one is true does the other have to be false)?

7. Does the evidence increase belief in more than one hypothesis at the same time?

8. Is there an alternative explanation for the evidence?

9. For each piece of evidence you wish to take into account, consider:

9.1. (How much) would you expect to see the observed evidence if the hypothesis were true?

9.2. (How much) would you expect to see the observed evidence if the hypothesis were false?

9.3. Is your expectation of seeing the evidence stronger if H is true or if H is false?

10. Are any pieces of evidence dependent on one another (i.e., is the probative value of one piece of evidence influenced by the existence of another)?

10.1. If so, what makes them dependent and how might this relationship affect their probative value?

10.2. Has statistical expert evidence been presented in the case? If so, are any of the factors or pieces of evidence relied on by the expert potentially dependent?

10.3. Have any pieces of evidence been left out due to irrelevance or deemed as too small in probative value? Why?

10.4. Is there any evidence that probative value may have been overestimated or underestimated? Why?

11. Does this piece of evidence relate to the hypothesis directly or through one or more inferential steps?

11.1. What assumptions are required to link this evidence to the hypothesis?

11.2. How strong or well-supported are these assumptions?

11.3. Is there an alternative explanation for each step in this chain?

APPENDIX B: PROPOSED INFORMATION SHEET

This information sheet is intended to assist you in interpreting and applying the checklist during reasoning and deliberation. The explanations below are illustrative only and are not intended to dictate conclusions.

Section 1: Framing the Hypotheses and Evidence

1. *What is the main hypothesis under consideration?*
2. *Have any alternative hypotheses been proposed? What are they?*
3. *What evidence is available in the case?*

Definitions:

Hypothesis = The main explanation, allegation or scenario being evaluated in light of the evidence.

Alternative hypotheses = Other plausible explanations that may also account for the events or evidence in the case.

Purpose:

These questions are intended to establish a clear structure for reasoning and case overview before the evaluation begins. They encourage the explicit identification of

what is being assessed, consider whether competing explanations exist and clarify what evidence is available in the case.

Probabilistic reasoning risks:

Errors may arise where the hypothesis or evidential basis remains implicit or insufficiently specified. Failure to explicitly identify competing hypotheses may result in premature closure or tunnel vision, where evidence is interpreted only in a manner consistent with a preferred explanation. There is also a risk of false dichotomy, where reasoning incorrectly assumes there are only two possible explanations while additional alternatives remain possible.

Example:

Main hypothesis = "The defendant is guilty of murdering X."

Alternative hypothesis = "X died of suicide"

Section 2: Prior Probabilities, Motive and Opportunity

4. Before considering any of the evidence, which hypothesis appears more probable?

4.1. Why do you consider this scenario more or less probable?

4.2. Is there a motive? If so, what impact does the presence of this factor have on the probability of the scenario(s)?

4.3. Is there opportunity? If so, what impact does the presence of this factor have on the probability of the scenario(s)

Definitions:

Prior probability = The probability of a hypothesis before considering any specific evidence in the case.

Purpose:

These questions encourage explicit reflection on the relative plausibility of competing explanations before evaluating the detailed evidence in the case. The framework does not require you to assign precise numerical probabilities. Rather, the purpose is to consider whether one explanation initially appears more probable than

another, and why. The questions concerning motive and opportunity are intended to assist you in identifying contextual factors that may influence the plausibility of competing hypotheses.

Probabilistic reasoning risks:

Failure to consider the prior probability of competing explanations constitutes a well-documented probabilistic fallacy commonly referred to as base-rate neglect. Base-rate neglect occurs where attention is focused primarily on striking or emotionally compelling evidence while insufficient consideration is given to how common or uncommon competing explanations are before considering any evidence.

Example:

Consider the hypotheses from above (H1 = Murder, H2 = Suicide). Before considering any of the evidence, H2 appears more probable, as suicide is more common than murder.

If the defendant has a motive to commit murder, e.g. an angry ex-partner, the probability of H1 increases.

If the defendant had opportunity, e.g. they were at the house when the murder was committed, the probability of H1 increases.

SECTION 3: COMPETING HYPOTHESES

5. Are the competing hypotheses exhaustive (i.e., do they cover all possible outcomes, meaning that at least one must be true)?

6. Are they mutually exclusive (i.e., if one is true does the other have to be false)?

7. Does the evidence increase belief in more than one hypothesis at the same time?

8. Is there an alternative explanation for the evidence?

Definitions:

Exhaustive hypotheses = Hypotheses that collectively cover all possible outcomes such that at least one must be true.

Mutually exclusive hypotheses = Hypotheses where, if one is true, the other must be false.

Purpose:

These questions encourage broader consideration of competing explanations and whether evidence may support more than one hypothesis simultaneously. They are intended to prevent reasoning from becoming restricted to a single preferred narrative.

Probabilistic reasoning risks:

These questions are intended to reduce the risk of false dichotomy. False dichotomy occurs where reasoning incorrectly assumes that competing hypotheses are both exhaustive and mutually exclusive. This may result in the mistaken belief that evidence supporting one explanation necessarily undermines all alternatives to the same degree.

Failure to consider alternative explanations may also encourage premature closure and overconfidence in one interpretation of the evidence.

Example:

Consider the hypotheses from above (H1 = Murder, H2 = Suicide). They are not exhaustive. The individual could have died from natural causes (e.g., a terminal illness). They are not exclusive, if H1 is false, i.e., no murder occurred, it does not mean that H2 is true, i.e., tha a suicide occurred.

SECTION 4: EVALUATING EVIDENCE AND INVERSION FALLACIES

9. For each piece of evidence you wish to take into account, consider:

9.1. (How much) would you expect to see the observed evidence if the hypothesis were true?

9.2. (How much) would you expect to see the observed evidence if the hypothesis were false?

9.3. Is your expectation of seeing the evidence stronger if H is true or if H is false?

Purpose:

These questions encourage you to evaluate evidence comparatively rather than in isolation.

The relevant issue is not simply whether evidence is consistent with a hypothesis, but whether the evidence is more expected if the hypothesis is true than if it is false.

These questions therefore encourage consideration of both:

— how well the hypothesis explains the evidence,

— and whether the same evidence could reasonably occur if the hypothesis were false.

Probabilistic reasoning risks:

These questions are intended to reduce the risk of inversion fallacies. An inversion fallacy occurs where the probability of observing the evidence if the hypothesis is true is incorrectly treated as equivalent to the probability that the hypothesis is true because the evidence exists.

This may lead judges to incorrectly assume that rare or unusual evidence necessarily implies a high probability of guilt. The relevant question is not whether the evidence appears incriminating in isolation, but whether the hypothesis provides a better explanation for the evidence than competing alternatives.

Example:

Consider the case above. Suppose DNA evidence is found matching the defendant in the victim's house. The relevant question is not whether the DNA evidence is consistent with the defendant's guilt, but whether it is more likely to see the DNA evidence if the defendant is guilty than innocent. Suppose the defendant was not known to the victim, the probability of seeing the DNA evidence at the crime scene is more likely if the defendant is guilty than innocent. Assuming the defendant also lives in the house, it is equally likely to find DNA evidence independent of whether they are innocent or guilty.

SECTION 5: CONVERGENT EVIDENCE, DEPENDENCIES AND PROBATIVE VALUE

10. Are any pieces of evidence dependent on one another (i.e., is the probative value of one piece of evidence influenced by the existence of another)?

10.1. If so, what makes them dependent and how might this relationship affect their probative value?

10.2. Has statistical expert evidence been presented in the case? If so, are any of the factors or pieces of evidence relied on by the expert potentially dependent?

10.3. Have any pieces of evidence been left out due to irrelevance or deemed as too small in probative value? Why?

10.4. Is there any evidence that probative value may have been overestimated or underestimated? Why?

Definitions:

Dependent evidence = Evidence where the probative value of one item is influenced by another.

Independent evidence = Evidence that retains its probative value regardless of the existence of other evidence.

Probative value = The extent to which evidence assists in proving or disproving a hypothesis.

Purpose:

These questions encourage reflection on how multiple pieces of evidence interact with one another. They are intended to assist judges in identifying whether multiple pieces of evidence genuinely provide separate support for a hypothesis or instead rely on overlapping assumptions, sources, or reasoning chains.

Probabilistic reasoning risks:

These questions are intended to reduce the risks of dependence neglect and convergence neglect. Dependence neglect occurs where dependent evidence is incorrectly treated as independent, artificially inflating the apparent strength of the case. Convergence neglect occurs where the combined strength of genuinely independent evidence is underestimated.

Statistical expert evidence may be particularly vulnerable to hidden assumptions of independence. There is also a risk that evidence may be overestimated or underestimated where its relationship to other evidence is not properly recognised.

Example:

Consider two witnesses who discussed the events together before giving their testimonies, these two witness statements are dependent, lowering their individual probative values. In the case of statistical evidence, forensic analyses may rely on the same underlying source or assumption. For example, in the Sally Clark case, the two infant deaths were initially treated as statistically independent events. However, if one child dies of SIDS, the probability of a second child in the same family also dying of SIDS may increase due to shared genetic or environmental factors. Treating

the two deaths as fully independent therefore risked overstating the improbability of the events occurring naturally.

SECTION 6: INFERENTIAL REASONING AND LINK-SKIPPING

11. Does this piece of evidence relate to the hypothesis directly or through one or more inferential steps?

11.1. What assumptions are required to link this evidence to the hypothesis?

11.2. How strong or well-supported are these assumptions?

11.3. Is there an alternative explanation for each step in this chain?

Definition:

Inferential step = A stage of reasoning required to connect evidence to a conclusion.

Assumption = A proposition that must be accepted in order for evidence to support a conclusion.

Chain of inference = A sequence of inferential steps connecting evidence to the hypothesis.

Purpose:

These questions encourage judges to examine how evidence supports a conclusion and whether intermediate assumptions have been made explicit. Some evidence supports a hypothesis directly, while other evidence supports a conclusion only indirectly through multiple inferential steps. These questions encourage reflection on whether each inferential step is adequately supported and whether alternative explanations remain possible.

Probabilistic reasoning risks:

hypothesis directly, while other evidence supports a conclusion only indirectly through multiple inferential steps. These questions encourage reflection on whether each inferential step is adequately supported and whether alternative explanations remain possible.

Probabilistic reasoning risks:

These questions are intended to reduce the risk of link-skipping. Link-skipping occurs where intermediate assumptions or inferential steps are ignored, causing the relationship between evidence and conclusion to appear stronger than it actually is. Failure to recognise inferential gaps may result in overestimation of the probative value of the evidence chain. The more inferential steps required, the greater the possibility that uncertainty or alternative explanations exist at each stage.

Example:

Suppose fibres matching the defendant's hoodie are found at the crime scene. The evidence does not directly establish guilt. Another person could have been wearing the hoodie, therefore this rests on the intermediate assumption: the defendant was the person wearing the hoodie at the crime scene.

Quaestio facti. Revista Internacional sobre Razonamiento Probatorio
Quaestio facti. International Journal on Evidential Legal Reasoning
Sección: Conjeturas y refutaciones
2026 I 11 pp. 225-232
Madrid, 2026
DOI: 10.33115/udg_bib/qf.i11.23292
Marcial Pons Ediciones Jurídicas y Sociales
© Bojan Spaić
ISSN: 2604-6202
Recibido: 28/05/2026 | Aceptado: 12/06/20266 | Publicado *online*: 30/06/2026
Editado bajo licencia Reconocimiento 4.0 Internacional de Creative Commons

WHAT GOES AROUND COMES AROUND: WHY THE CRITIQUE OF EXPERT-LED INTEGRAL BAYESIAN MODELLING ALSO UNDERMINES JUDGE-LED BAYESIAN MODELLING*

Bojan Spaić
University of Belgrade Faculty of Law

ABSTRACT: This article offers an internal refutation of Anne Ruth Mackor's defence of court-led integral Bayesian modelling in criminal adjudication. Mackor usefully distinguishes between the now well-established Bayesian treatment of individual items of forensic evidence and the more controversial Bayesian modelling of criminal cases as a whole. She also convincingly argues that expert-led whole-case modelling is objectionable because the expert must select hypotheses, evidence, dependencies, and probabilistic inputs in a manner that effectively places the expert in the judge's chair. Yet she maintains that this objection applies with less force when whole-case Bayesian modelling is performed by judges themselves, especially as an internal "means of inspection" or "sharpening" device and with the support of a forensic adviser. This article argues that the distinction fails. The decisive objection to integral Bayesian modelling does not arise merely from the formal identity of the modeller. It lies in the modelling choices themselves being adjudicatively decisive. Those choices remain no less decisive when made by judges, and the involvement of the forensic adviser reintroduces, in a less visible form, the very institutional concern Mackor's critique seeks to avoid.

KEYWORDS: evidential reasoning; legal probabilism; legal abductivism; Anne Ruth Mackor.

* This paper is a contribution to the 2026 annual strategic project of the Faculty of Law University of Belgrade "Problems of creation, interpretation and application of law."

SUMMARY: 1. INTRODUCTION.— 2. STRENGTHS AND WEAKNESSES OF JUDICIAL BAYESIANISM.— 3. WHY THE SAME OBJECTION APPLIES TO JUDGE-LED MODELLING.— 4. THE FORENSIC ADVISER AS HIDDEN CO-AUTHOR.— 5. THE INSTABILITY OF THE "MEANS OF INSPECTION".— 6. CONCLUSION.— LITERATURE

1. INTRODUCTION

Anne Ruth Mackor's article on Bayesian modelling of criminal cases as a whole is a careful intervention in a debate that is often marred by needless disagreement and conceptual confusion. Its merit primarily lies in its insistence that one must distinguish between different uses of Bayesian reasoning in criminal adjudication. Mackor emphasises that Bayesian methods are widely accepted for the likelihood-ratio assessment of individual items of forensic evidence, and that the genuinely controversial question concerns the integral modelling of the criminal case as a whole. She therefore rightly criticises judicial reactions that blur the difference between objections to one particular use of Bayes and objections to Bayesian reasoning as such (Vlek *et al.*, 2016, p. 285).

The article's central claim, however, is unstable. Mackor argues that Dutch lower courts were right to reject expert-led integral Bayesian modelling because such modelling requires the expert to make a series of choices concerning hypotheses, evidence, interdependencies, and probabilistic inputs that are not merely technical but adjudicatively decisive. On her account, the problem with the expert-led model is not that the Bayesian method is itself unsound. The problem is that in constructing the model, the expert effectively serves as the judge. (Mackor, 2026, pp. 364-365) The expert selects, structures, and evaluates. That, Mackor claims, is institutionally objectionable from a rule-of-law perspective.

Mackor then takes a judicial turn. She contends that the same objection does not hold, or does not hold in the same way, where judges themselves undertake integral Bayesian modelling, especially in the deliberation chamber, and especially where such modelling functions as an internal "means of inspection" after a more traditional evidential assessment has already been formed. She goes on to develop this possibility in more concrete terms by proposing that judges may, with the support of a forensic adviser, construct Bayesian networks to compare their traditional judgment with the outcome of their own Bayesian analysis, thereby sharpening their evidential decision (Mackor, 2026, pp. 371-372).

This article argues that Mackor's move from rejecting expert-led integral Bayesian modelling to defending judge-led integral Bayesian modelling does not succeed. The objections she raises against expert-led modelling cannot be limited to the expert-led version alone, because the very choices that make expert-led modelling objectionable remain present and decisive in judicial modelling as well. If the vice of expert-led modelling is that the evidential architecture of the case is structured through contest-

able and outcome-sensitive choices, then changing the identity of the modeller does not by itself neutralise that vice. The problem lies not simply in who models the case, but in what it means to model a criminal case as a whole.

The argument proceeds in four steps. (1) I reconstruct the relevant points of Mackor's position. (2) I show that the objection she raises against expert-led integral modelling applies, mutatis mutandis, to judge-led modelling as well. (3) I then argue that the role she reserves for the forensic adviser reintroduces, in internalised form, the very problem of expert displacement that grounds her critique of the expert-led model. (4) Finally, I argue that her proposal to treat judicial Bayesian modelling as a "means of inspection" or "double-check" is methodologically unstable. My claim is modest. I do not argue that Bayesian reasoning is always inappropriate in law, nor that integral Bayesian modelling could never, under any circumstances, prove superior to its competitors. I argue only that Mackor's own premises do not justify the distinction on which her judicial alternative depends (Di Bello, 2019).

2. STRENGTHS AND WEAKNESSES OF JUDICIAL BAYESIANISM

An important aspect of Mackor's paper is to be found in her refusal to attack Bayesian methods indiscriminately. She repeatedly stresses that the Bayesian model is widely accepted as a method for estimating the likelihood ratio or probative force of forensic evidence, and that Dutch case law eventually clarified that the real controversy begins only when Bayesian reasoning is extended from individual forensic evidence to the modelling of a criminal case as a whole. This distinction is correct and valuable. It prevents a familiar overreaction in which scepticism about whole-case probabilistic modelling is transformed into scepticism about all Bayesian reasoning in adjudication (Vlek *et al.*, 2016, p. 288).

Even more important is Mackor's diagnosis of what is actually wrong with expert-led integral modelling. She does not say, at least in her own voice, that Bayes is too mathematical, too foreign to adjudication, or too controversial simply because probabilities are involved. On the contrary, she explicitly argues that "the method itself is not disputed among forensic experts" and that the real difficulty arises in "the selection and evaluation of the hypotheses, the evidence and their interdependencies." (Mackor, 2026, p. 371) On that basis, she explains why an expert who undertakes an integral analysis of a criminal case as a whole risks taking the place of the court: because the expert's selection and evaluation will not necessarily correspond to that of the judge and may therefore amount to placing "the expert on the court's chair."

This is one of the main points in the article. Whole-case Bayesian modelling is an exercise in constructing the evidential architecture of the case. Someone must formulate the hypotheses, determine which evidence matters, decide how pieces of evidence are related, and assign or approximate the relevant probabilistic force (Hunt & Mostyn, 2020). Those decisions are not neutral. They shape the path to the result.

Mackor is therefore right to see the expert-led model not as a harmless aid but as a possible displacement of adjudicative responsibility.

3. WHY THE SAME OBJECTION APPLIES TO JUDGE-LED MODELLING

When the strongest points of Macor's paper are taken seriously, the weakness of her alternative becomes apparent. The objection to expert-led modelling is that whole-case Bayesian analysis requires the making of adjudicatively decisive choices. But those choices do not disappear when judges themselves perform the modelling. They remain exactly the same choices: the formulation of hypotheses, the inclusion and exclusion of evidential items, the characterisation of dependence or independence, and the estimation or positing of priors and likelihoods where statistical grounding is weak or absent.[1] The only thing that changes is the institutional identity of the person who makes them.

This difficulty is central. Either those modelling choices are adjudicatively decisive, or they are not. If they are not, then Mackor's institutional objection to expert-led modelling is significantly weakened. If they are, judicial ownership cannot, by itself, answer the objection. The modelling process still structures the evidential field in a way that bears on the outcome. A change in authorship does not alter the character of the act (Allen, 2019, p. 15).

Mackor's own account makes this plain. She acknowledges that judges are not experts in Bayesian modelling and notes empirical work suggesting that judges often misunderstand Bayesian reports. She also emphasises that, outside well-supported forensic contexts, priors and many non-forensic evidential estimates cannot be grounded in reliable statistics and must therefore be estimated by the court. In other words, the judicial version retains the modelling choices that troubled Mackor in the expert-led context, and assigns them to actors who, by her own account, are often technically ill-equipped to manage them without assistance and must frequently do so on the basis of contestable, non-statistical judgments (Mackor, 2026, pp. 369-370).

The problem is therefore structural. Whole-case Bayesian modelling is a constructive activity. The case must be represented under descriptions. Some propositions must be elevated to the status of hypotheses, others to evidence, and others may drop out of the model entirely. Dependencies must be posited at one level rather than another. Some uncertainty must be stabilised into qualitative or numerical ranges (Pardo & Allen, 2008). None of this becomes less contestable merely because the modeller is now a judge rather than an external expert.

[1] Authors who argue in favor of Bayesian networks modeling emphasize the characteric dificulty of the procedure: "Representing a scenario is nonetheless still a difficult task due to all the probabilities that need to be elicited" (Vlek *et al.*, 2016, p. 318)

Indeed, one might say that Mackor's own reasoning generates a dilemma. If the decisive concern about expert-led modelling is that it structures the adjudicative field through contestable choices, then judge-led modelling inherits that concern because it performs the same structuring operation. If, on the other hand, judicial authorship is said to cure the problem simply because judges are entitled to make adjudicative choices, then Mackor's critique of expert-led modelling becomes oddly incomplete, since the mere existence of contestable modelling choices can no longer serve as the decisive reason for resistance. In that case, the problem would have to be redescribed more narrowly as one of institutional authority alone. But Mackor's own article says more than that. It emphasises not only who decides but also the fragility, discretionary character, and outcome sensitivity of what must be decided.

For that reason, the judicial alternative cannot be vindicated simply by insisting that judges, unlike experts, are entitled to decide cases. Of course they are. The question is whether modelling the case as a whole in Bayesian form introduces a set of structurally problematic operations into adjudication that judicial entitlement alone cannot purify.

4. THE FORENSIC ADVISER AS HIDDEN CO-AUTHOR

The instability of Mackor's position becomes clearer once the forensic adviser enters the scene. Mackor expressly envisages the possibility of judicial Bayesian modelling with the assistance of a forensic adviser. At one point, she notes that judges may be assisted in Bayesian modelling by the court's forensic adviser, while conceding that even then, "there is a risk that the forensic adviser's modelling proposals will exert undue influence on the judge's evidential decision" {Mackor, 2026 #36748@369}. Later, when describing her group's research agenda, she states that the relevant analysis should involve judges setting up a Bayesian network "with the help of a forensic advisor" {Mackor, 2026 #36748@378}.

This concession is more serious than the article allows. The forensic adviser is introduced because judges are not Bayesian experts and because modelling a case as a whole is technically and conceptually demanding. But the more necessary the adviser becomes, the more the original institutional objection returns. If expert-led modelling was objectionable because the expert shapes the evidential architecture of the case (Tribe, 1971, pp. 1381-1386), it is hard to see why the forensic adviser's analogous shaping role becomes innocuous merely because it is exercised inside the judicial process and under the label of assistance.

The adviser may not be the formal author of the final judgment, but he may still be a co-author of the model that structures the case. He may influence which nodes are created, how they are related, how alternative hypotheses are formulated, whether dependency is modelled at one level or another, and how verbal estimates are stabilised into ranges or numbers. Those are not peripheral matters. They are precisely

the matters Mackor had identified as decisive in explaining why expert-led integral Bayesian modelling threatens to put the expert in the judge's chair.

Nor is it enough to say that the judge remains free to reject the adviser's proposals. The same can be said of external expert reports in general. A court is always formally free to reject them. Mackor's objection to the expert-led model, however, was not defeated by that formal freedom. It was that the expert's modelling choices shape the evidential field itself. That point applies equally here. If the adviser is involved in building the network, the displacement concern has not disappeared but has been internalised.

There is, in fact, a further difficulty. In the expert-led model, the court and the parties at least encounter the model as the work of a separate actor and can criticise it as such. In the judicial version, by contrast, the adviser's technical influence risks becoming less visible as it is absorbed into the court's own deliberative process. The danger is not that the adviser openly replaces the judge. It is that expert influence becomes harder to identify because it is folded into what now appears as judicial self-reflection. The expert is no longer visibly in the judge's chair. He is standing behind it, arranging the model through which the court comes to understand its own reasoning.

5. THE INSTABILITY OF THE "MEANS OF INSPECTION"

Mackor's most sophisticated response to these worries is her suggestion that judge-led whole-case Bayesian modelling need not function as the primary basis of adjudication. She proposes, instead, that judges may first reach a preliminary evidential assessment in a "traditional" manner and then use integral Bayesian analysis as a "means of inspection," an internal check, or a "sharpening function." {Mackor, 2026 #36748@371}. Later, she develops this further by describing her group's investigation into judges who compare their traditional judgments with the results of their own Bayesian analyses {Mackor, 2026 #36748@377-381}.

If the Bayesian model merely confirms the traditional assessment, then its independent normative significance is limited. It adds discipline, perhaps, or rhetorical reassurance, but not much more. Yet Mackor wants it to reveal hidden assumptions, expose errors in reasoning, force attention to dependence relations and evidential levels, and improve the quality of judicial evidence evaluation {Mackor, 2026 #36748@367}.

But if the model can do that meaningfully, then it is not merely a check. It becomes capable of revising, unsettling, or competing with the traditional assessment. Once the Bayesian model acquires that capacity, the relation between the first-stage assessment and the second-stage model becomes methodologically unclear. If the two converge, the second risks redundancy. If they diverge, which governs? Mackor suggests that divergence should provoke critical comparison, but that is not yet a cri-

terion of resolution. A tool that is serious enough to change the court's view cannot be described as merely auxiliary without further explanation.

The difficulty becomes even sharper when Mackor turns to Bayesian networks. She stresses that the primary value of constructing a network lies not in the final numerical output but in the joint activity of structuring, discussing, and refining assumptions {Mackor, 2026 #36748@378}. At the same time, however, she notes that software can perform highly complex calculations once priors and evidential values are entered, and she defends quantification on the grounds that it can help judges grasp the cumulative effect of multiple likelihood ratios and priors. She also argues that quantification may encourage greater uniformity in the use of verbal probability judgments {Mackor, 2026 #36748@381}.

These are not negligible considerations. But they do not resolve the deeper instability. The more operational the model becomes, the more it threatens to become a rival decision framework rather than a mere reflective aid (Tribe, 1971, p. 1971). The sharper the "double-check," the less plausible it is to insist that it is only a check. Mackor's proposal thus oscillates between two poles. Either judicial Bayesian modelling is weak enough to remain subordinate, in which case its added value is modest (Hunt & Mostyn, 2020, p. 16). Or it is strong enough to alter the court's evidential outlook, in which case it assumes a decisional significance that her institutional framework does not adequately contain.

6. CONCLUSION

Mackor's article is at its best when it insists that the principal difficulty of whole-case Bayesian modelling lies not in Bayes as such, but in the construction of a probabilistic representation of the criminal case as a whole. Her most persuasive point is that expert-led integral modelling is objectionable because the modeller must select hypotheses, evidence, dependencies, and probabilistic inputs in a manner that risks displacing judicial responsibility. That insight is both important and correct.

But the insight cannot be confined to the expert-led model alone. The same modelling choices remain necessary when judges themselves undertake integral Bayesian modelling. Those choices do not become less contestable, less theory-laden, or less outcome-sensitive merely because their author is now the court. Nor does the involvement of the forensic adviser neutralise the problem. It reintroduces, in a less visible form, the very concern about expert influence that grounds Mackor's original critique. Finally, the attempt to stabilise the judicial alternative by redescribing it as a "means of inspection" or "double-check" fails to dispel the difficulty. The model is either too weak to matter greatly or strong enough to become a rival framework of adjudication.

The present argument is deliberately limited. It does not deny that Bayesian methods may have an important place in criminal adjudication (Di Bello, 2019, p. 6). It

does not deny that integral Bayesian modelling might, under some conditions, prove superior to its alternatives. It argues only that Mackor has not shown that judge-led integral Bayesian modelling escapes the very objection on which her critique of expert-led modelling depends. On her own premises, the objection returns. The same objection arises twice.

LITERATURE

Allen, R. J. (2019). Standards of Proof and The Limits of Legal Analysis. *Diritto & questioni pubbliche*, *XIX*, 7–25.

Di Bello, M. (2019). Plausibility and probability in juridical proof. *The International Journal of Evidence & Proof*, *23*(1-2), 161–167. *https://doi.org/10.1177/1365712718815355*

Hunt, I., & Mostyn, J. (2020). Probability reasoning in judicial fact-finding. *The International Journal of Evidence & Proof*, *24*(1), 75–94. *https://doi.org/10.1177/1365712719875753*

Mackor, A. R. (2026). Bayesian Modelling of Criminal Cases as a Whole. *Quaestio facti. Revista Internacional sobre Razonamiento Probatorio*, *10 Quaestio facti. International Journal on Evidential Legal Reasoning*, 363–384. *https://doi.org/10.33115/udg_bib/qf.i10.23299*

Pardo, M. S., & Allen, R. J. (2008). Juridical Proof and the Best Explanation. *Law and Philosophy*, *27*(3), 223–268.

Tribe, L. H. (1971). Trial by Mathematics: Precision and Ritual in the Legal Process. *Harvard Law Review*, *84*(6), 1329. *https://doi.org/10.2307/1339610*

Vlek, C. S., Prakken, H., Renooij, S., & Verheij, B. (2016). A method for explaining Bayesian networks for legal evidence with scenarios. *Artificial Intelligence and Law*, *24*(3), 285–324. *https://doi.org/10.1007/s10506-016-9183-4*

www.quaestiofacti.com