Quaestio facti

Revista Internacional sobre Razonamiento Probatorio
International Journal on Evidential Legal Reasoning

2026 | N. 10

ISSN: 2604-6202

ENSAYOS

HOMENAJE A LARRY LAUDAN

Rachel Laudan, *Behind the CV: Reflections on Larry Laudan's Unconventional Academic Career*

Amalia Amaya y Andrés Páez, *El legado de Larry Laudan y el futuro de la epistemología jurídica*

Ernesto Saccardi, *Laudan's Meta-Rule for Evaluating Evidence and the Case of Expert Testimony*

Andrés Páez, *La negligencia institucional frente al riesgo epistémico de los sesgos cognitivos en el derecho*

Rodrigo Coloma Correa, Renato Lira Rodríguez y Sebastián Sandoval Ayala, *Dudas razonables. Sobre la concepción de estándar de prueba de Larry Laudan*

Hylke Jellema, *Standards of Proof for Future Crimes and Decision Theory*

Edgar Aguilera, *¿Fábrica de inocentes? ¡No tan deprisa! (De un posible sesgo en el proyecto de epistemología jurídica de Laudan)*

Eunseol Lee, *The Pursuit of Truth and the Justification in Legal Fact-Finding: A Comparative Analysis of Bentham, Laudan, and Damaška*

Michael S. Pardo, *Legal Epistemology and Legal Proof*

Sebastián Sandoval Ayala, *La prueba de los hechos futuros. Consideraciones semánticas y epistémicas para la toma de decisiones*

Mauricio Duce J., *Los informes periciales escritos en el proceso penal en Chile: análisis crítico de su regulación y práctica*

Sergio Cruz Arenhart, *La conducta de las partes como prueba en el proceso civil brasileño*

Alejandra Olvera Dorantes, *Motivos para repensar la violencia política contra las mujeres por razones de género en la legislación mexicana: un enfoque desde el razonamiento probatorio*

CONJETURAS Y REFUTACIONES

Anne Ruth Mackor, *Bayesian Modelling of Criminal Cases as a Whole a Philosophical Reflection on Dutch Case Law*

Càtedra de Cultura Jurídica | Universitat de Girona

Marcial Pons

Quaestio *facti*

Revista Internacional sobre Razonamiento Probatorio
International Journal on Evidential Legal Reasoning

2026 | N. 10

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

Jordi Ferrer Beltrán
Universitat de Girona, España

COMITÉ DE REDACCIÓN

Daniela Accatino
Universidad Austral de Chile, Chile

Daniel González Lagier
Universidad de Alicante, España

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

Carmen Vázquez
Universitat de Girona, España

SECRETARÍA DE REDACCIÓN Y EDICIÓN

Anna de Castro Cros
Universitat de Girona, España

COMITÉ EDITORIAL

Flavia Carbonell
Universidad de Chile, Chile

Antonio Manzanero
Universidad Complutense de Madrid, España

Mauricio Duce
Universidad Diego Portales, Chile

Florencia Rimoldi
Universidad de Buenos Aires, Argentina

Mercedes Fernández
Universidad de Alicante, España

Sarah Summers
University of Zurich, Suiza

Raymundo Gama
Instituto Tecnológico Autónomo de México

Giovanni Tuzet
Università Bocconi, Italia

Ho Hock Lai
National University of Singapore, Singapur

Jonatan Valenzuela
Universidad de Chile, Chile

CONSEJO ASESOR

Ronald Allen
Northwestern University, EEUU

Amalia Amaya
Universidad Nacional Autónoma de México, México

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

Marina Gascón
Universidad de Castilla-La Mancha, España

Susan Haack
University of Miami, EEUU

Rachel Herdy
Universidad Adolfo Ibáñez, Chile

John Jackson
University of Nottingham, Reino Unido

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

Catherine Piché
Université de Montréal, Canadá

Joan Picó i Junoy
Universitat Pompeu Fabra, España

Giovanni Priori
Pontificia Universidad Católica del Perú, Perú

Vitor Lia de Paula Ramos
Pontificia Universidad Católica do Rio Grande do Sul, 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

ÍNDICE

HOMENAJE A LARRY LAUDAN
(editado por Amalia Amaya y Andrés Páez)

Rachel Laudan, *Behind the CV: Reflections on Larry Laudan's Unconventional Academic Career* 11

Amalia Amaya y Andrés Páez, *El legado de Larry Laudan y el futuro de la epistemología jurídica* 25

Ernesto Saccardi, *Laudan's Meta-Rule for Evaluating Evidence and the Case of Expert Testimony* 29

Andrés Páez, *La negligencia institucional frente al riesgo epistémico de los sesgos cognitivos en el derecho* 47

Rodrigo Coloma Correa, Renato Lira Rodríguez y Sebastián Sandoval Ayala, *Dudas razonables. Sobre la concepción de estándar de prueba de Larry Laudan* 67

Hylke Jellema, *Standards of Proof for Future Crimes and Decision Theory* 95

Edgar Aguilera, *¿Fábrica de inocentes? ¡No tan deprisa! (De un posible sesgo en el proyecto de epistemología jurídica de Laudan)* 125

Eunseol Lee, *The Pursuit of Truth and the Justification in Legal Fact-Finding: A Comparative Analysis of Bentham, Laudan, and Damaška* 157

Michael S. Pardo, *Legal Epistemology and Legal Proof* 187

ENSAYOS

Sebastián Sandoval Ayala, *La prueba de los hechos futuros. Consideraciones semánticas y epistémicas para la toma de decisiones* 223

Mauricio Duce J., *Los informes periciales escritos en el proceso penal en Chile: análisis crítico de su regulación y práctica* 245

Sergio Cruz Arenhart, *La conducta de las partes como prueba en el proceso civil brasileño* 283

Alejandra Olvera Dorantes, *Motivos para repensar la violencia política contra las mujeres por razones de género en la legislación mexicana: un enfoque desde el razonamiento probatorio* 327

CONJETURAS Y REFUTACIONES

Anne Ruth Mackor, *Bayesian Modelling of Criminal Cases as a Whole a Philosophical Reflection on Dutch Case Law*..... 363

HOMENAJE
A LARRY LAUDAN

(editado por Amalia Amaya y Andrés Páez)

Quaestio facti. Revista Internacional sobre Razonamiento Probatorio
Quaestio facti. International Journal on Evidential Legal Reasoning
Sección: Homenaje a Larry Laudan
2026 I 10 pp. 11-24
Madrid, 2025
Marcial Pons Ediciones Jurídicas y Sociales
© Rachel Laudan
ISSN: 2604-6202
Editado bajo licencia Reconocimiento 4.0 Internacional de Creative Commons

BEHIND THE CV: REFLECTIONS ON LARRY LAUDAN'S UNCONVENTIONAL ACADEMIC CAREER

Rachel Laudan

Larry's academic career, as is typical, is summed up by that most ritualized of documents, the curriculum vitae, a dry catalog of education, awards, publications, and academic positions. As far as the latter goes, his cv shows that at the age of 40 Larry resigned from a top flight philosophy program at the University of Pittsburgh, one that he had helped build. He moved to a fledgling science studies center at Virginia Tech, a university strongly oriented to agriculture and engineering, and consequently weak in the humanities. Shortly thereafter, Larry took a position as chair of the philosophy department at the University of Hawaii, an institution marginal in every way to American academic philosophy. In his mid 50s, he left academia. At 60 he rejoined it, accepting a position at the National University of Mexico, pivotal in the Spanish-speaking world, but unknown to most American academics. This succession of university appointments was characterized by a friend as "scarcely conventional" and regarded by most of his colleagues as something between quixotic and disastrous.

This "scarcely conventional" career resulted from Larry's effort to reconcile his lifelong pursuit of an intellectual problem with my own academic career, against the background of academic politics. The intellectual problem was that of understanding how to improve the inferences from the evidence to justified belief whether in a scientific theory, a legal verdict, or in life more generally. My career, like all women's in the 1970s, had to be forged in the context of the century-long reluctant opening of academia to others than white protestant males—Jews, working class, women, and blacks among other. The career was further complicated by my penchant for subjects unfashionable in my discipline of history and philosophy of science: geology,

technology, and food. For both of us, the relevant academic politics were the rigid university pecking order and the departmental micropolitics surrounding positions and research fashions.

For all this, Larry ended up feeling that his career, for all its ups and downs, had been deeply satisfying (as I did too). So this essay is less a defense of his decisions, than an invitation to new generations of scholars who, in the United States and elsewhere, face universities struggling to define their place in difficult political times, to consider careers other than the prescribed conventional ones. They would not, of course, take the same shape as Larry's as times are different, and so are the alternatives that might be possible.

I shall say nothing about Larry in the classroom although he was a born teacher. Nor shall I focus on his administrative work, though understanding that without institution building his intellectual ambitions would come to nothing, he handled work as chair, editor, conference organizer with skill and dispatch.

LARRY'S EARLY CAREER

Larry's early academic career in the 1960s, a time when American universities were expanding, was unusual only in being particularly fast track. From his early teens; Larry had been fascinated by astrophysics and the way in which small data points could be used to construct sweeping theories of the universe. Thus, after high school in Kansas City, he majored in physics at the local state university, the University of Kansas. From there, with his first wife, he went to Princeton. His goal was an astrophysics Ph.D. directed by the theoretical physicist, John Wheeler.

Discovering that his dissertation would depend on mathematical prowess, not on probing why it was more reasonable to believe one theory of the universe rather than another, he looked for a more congenial program. History and philosophy of science, just then being established in American universities with support from the National Science Foundation, promised a different way of understanding when theories were worthy of acceptance. He began his thesis on *The Idea of a Physical Theory from Galileo to Newton: Studies in 17th-Century Methodology* with a committee consisting of the distinguished philosopher of science Carl G. (Peter) Hempel famed for his work on explanation in science, Thomas Kuhn who had just published *The Structure of Scientific Revolutions*, and Charles Gillispie, a leading historian of science.

Larry found the twinned histories and philosophies of science by the Victorian scientific polymath, William Whewell and the French physicist, Pierre Duhem particularly inspiring. They had turned to history to elicit how science had progressed by choosing successively better theories, thus surging ahead of the humanities in terms of reliable knowledge. With those models in mind, Larry explored how leading scientists in the seventeenth-century scientific revolution justified their choices of new

theories. Like his heroes, in the back of his mind was the hope that this learning how to learn could be applied more generally outside the sciences.

Larry's committee was on the point of recommending him for positions in prestigious American universities when Larry announced that he was going to work with the pioneering historian of philosophy of science, Gerd Buchdahl in Cambridge, England. Cambridge too was an elite university, but it was outside the sphere of influence of American academics. Since the hierarchy of American universities depended on faculty at the elite universities placing their graduate students at closely similar institutions, they were miffed. He was told he could no longer count on their support.

In England, Larry completed his dissertation. He accepted a lectureship (assistant professorship) at another major institution, University College London, in the world's oldest history and philosophy of science department, perfect for his interests. At the same time, he, with Gerd Buchdahl, founded and edited *Studies in the History and Philosophy of Science* to create a place to publish the longform articles that they believed essential to an integrated history and philosophy of science.

Larry by now had carved out his intellectual niche. Neither at that time nor later did Larry see himself primarily as a philosopher. While he admired the clarity and argumentative precision of the analytic philosophy entrenched in Anglophone philosophy departments, he had no interest in metaphysics or traditional epistemology, let alone ethics or aesthetics. More important, he rejected the assumption that analysis of language and concepts would yield worthwhile knowledge, progress of knowledge, well-founded beliefs.

Instead philosophy, like science, was an empirical enterprise. His particular interest in learning how to improve inference from data to belief could not be answered without resort to particular, comparative cases, and thus to history. To clarify his ideas, Larry scrutinized the small number of crucial papers and then turned to conversation. For him, the back and forth of a graduate seminar, the question and answer following a talk, an informal conversation, or an internal dialogue were not only fun but intellectually crucial. (For those who did not know him well, the intensity of these conversations could be daunting). This was no ivory tower exercise. The results had practical importance, part of the broader project of how human decision making could be improved by clarifying the relations between evidence and belief.

In 1969, the philosopher of science Adolf Grunbaum offered Larry a position in the University of Pittsburgh Philosophy Department. Grunbaum was building the department into one of the two or three best in the nation, so even though the University as a whole was only mid-ranked, this was an appealing offer, especially with a large salary increase. Given Larry's interests, Grunbuam arranged for a joint appointment in History.

On arriving in Pittsburgh, Larry discovered that, because of departmental micropolitics, the situation did not live up to his expectations. Mainstream historians

regarded history of science not as the key to progress in knowledge, but as marginal and elitist compared to labor and environmental history. That administrators wanted the newly-available funding from the National Science Foundation that historians of science could bring in, only made the resentment worse. Larry was cold shouldered.

Larry thus enlisted Grunbaum's help in founding an independent Department of History and Philosophy of Science. Given the institutional setting, his interests and the department's tilted in the philosophical direction. By 1972, Larry was Chair and Full Professor. He was 31 years old. He probaby assumed he would spend the rest of his career building the department, and carrying on those crucial conversations with colleagues and graduate students and with visiting scholars from across the States and beyond that arrived every year.

OUR TWO-CAREER PROBLEM

But no. In 1972, Larry and I ran into each other at a meeting of the American Academy for the Advancement of Science in Washington, D.C. We had met some years earlier at University College London when I was a graduate student in History and Philosophy of Science, but both married, we had had only the most formal of relations. Shortly after this second meeting, we both left our first marriages, a decision that was made swiftly, but not impulsively. We were married two years later in 1974.

Unusually for the 1950s, my intellectual family had raised me to expect a career, not simply a job to supplement the family income. My all girls 'public' school, staffed by first-wave-feminist women, reinforced that expectation. Consequently, as an undergraduate I happily embarked on a degree in geology, a field that was hardly hospitable to women and won a First Class Honors degree. I was accepted to pursue a Ph.D. at the premier Lamont-Doherty Observatory at Columbia University. Before leaving, I stumbled onto the works of Stephen Toulmin and June Goodfield, Norwood Russell Hanson, and Thomas Kuhn in the university library. Enthralled by their revelatory understanding that the history of science was not just a random sequence of inventions and discoveries but linked by underlying intellectual traditions, I instead signed up for a History and Philosophy of Science Ph.D. at University College London.

Sad to say, the program was less than ideal. As was common in the British system, it involved almost no coursework or supervision, so what I knew of history came from my high school days and I learned nothing of philosophy. On my own, I decided to explore the links between the Industrial Revolution, the emergence of geology as a science, and the method of geological mapping. Maps were to geology what mathematics was to physics and astronomy, I decided, ways of concisely representing the phenomena and of making surprising predictions.

On the personal side, what would now be called rampant harassment was the order of the day at both University College London and the associated London School of Economics. A network that stretched from those institutions to universities on the east and west coasts of the United States took it for granted that any woman pursuing a Ph.D. would go to bed with them. Ironically this was an unintended consequence of the second feminist wave which was arguing for sexual liberation. Many male faculty considered seducing female students was the moral high road to the desired aim of destroying traditional marriage. It was not peculiar to those institutions. I was thoroughly confused. My conversations with other women academics in the last year or two made it clear how many women in Anglophone abandoned career hopes rather than be subject to it. I survived. Most didn't. "Why had I not reported this at the time?" asked Larry later. "Apart from having no idea to whom I would report it, would they have believed what I said about internationally-known full professors? Would you?" "No," he said. Point made.

In spite of all this, I realize in retrospect that I worked my way toward my own intellectual niche, which turned out to be distinct from Larry's even though our enthusiasm for history and philosophy of science was mutual. I was fascinated by uncovering the intellectual traditions that bind seemingly disparate historical events together. Where Larry leaned on conversation to clarify his ideas, I loved documents—books, maps or archives. On doing research relevant to an audience wider than just academics, we converged again. These insights came years later though. At the time I was glad to find the English equivalent to an American tenure-track position.

On arriving in Pittsburgh, a two-career academic marriage looked possible. I was offered a tenure-track position in History at Carnegie-Mellon University, an excellent institution next door to the University of Pittsburgh, to teach history of technology to engineering students. Luckily the hiring committee was unaware that history and philosophy of science assiduously avoided any connection with grubby history of technology. And because my dissertation dealt with the practical as well as the theoretical side of geology, I had learned a good bit of history of technology on my own.

Then in my third year, a new Dean of Arts and Humanities needed positions for a new Social Sciences Department. I, along with the rest of the untenured faculty was fired. Larry now had to face up to the difficult decisions a two-career marriage could entail. In spite of our premarital conversations, I think he had always assumed that, like his mother and first wife, I would be content to work at something pleasant to supplement the family income, but would not take on anything that would clash with his career.

We began considering the options. A tenure-track position at the University of Pittsburgh was out of the question. Institutional anti-nepotism policies were in place that (largely rightly) did not allow hiring or promotion of family members. Even had this not been the case, Philosophy and History and Philosophy of Science, like most other American academic departments, were overwhelmingly male, at best passively

resistant, at worst openly hostile, to women as colleagues. And it was obvious I was ignorant of philosophy. This was compounded by my research in geology. Philosophers of science, Larry included, took it for granted that theoretical physics was the archetypal science, that biology was barely worth consideration except insofar as it could be derived from physics, and that geology, lacking mathematical expression and universal laws (so they assumed) was not even worthy of the name science. Finally, although I was only two years younger than Larry, thanks to a "gap year" teaching in Nigeria and other interruptions, I was a generation behind him academically. The career gap made it all too easy for me to be characterized as an unserious, unscrupulous bimbo.

Maybe I should take a job in one of the many corporations headquartered in Pittsburgh? Maybe I should pursue a career in law? No. Maybe I should commute to a university in another city? No again. By now I knew Larry would not flourish with a wife in a nine to five job, let alone another city. He needed support and understanding—whether for his eccentric working hours from noon to 3 am, or his constant pipe smoking, or even for the endless dinners and cocktail parties his work involved, not to mention reading and copy editing drafts of publications.

I took an adjunct position at the University of Pittsburgh and did the day-to-day administration of the Philosophy of Science Center. Both were made-up positions to accommodate a spouse, the kind of positions that were to become increasingly common in American academia. While I was initially grateful for the efforts that had been made to find some kind of place for me, I grew increasingly resentful that these were dead end positions. Meanwhile Larry continued to publish, including his well-received first book, *Progress and Its Problems* (1977) and took on the Directorship of the Philosophy of Science Center.

After several years, I explained to Larry, rather emotionally I remember, that these stop gaps were not working. Our increasingly unequal status was wreaking havoc with the kind of partnership we had envisioned. Since Larry's children, who meant the world to him, now lived far away, visiting only a couple of times a year, that formerly compelling reason to stay in Pittsburgh had vanished. Perhaps the time had come to make good on his promise that I could pursue a career and look for two positions in one place. We talked this over with the only academic couple we knew in Pittsburgh, Rolland and Cristina Paulsten (Education and Linguistics respectively). With their support, we decided to go on the job market.

THE MOVE TO A LESS-PRESTIGIOUS DEPARTMENT

In 1980, the Dean of Science and Humanities at Virginia Tech offered Larry tenure in Philosophy, me my desired tenure-track position, in History as seemed appropriate, with joint appointments in the new Science Studies Center being set up under the leadership of the historian of science, Arthur Donovan for both of us. In

considering whether to accept, I knew I was asking a lot of Larry. Although the salary offered was good, Larry knew he would miss his valued colleagues and graduate students, perhaps have a higher teaching load and less money for travel and sabbaticals, and a lesser library. He feared he would lose his place at the center of the national network of scholars working on related problems.

And all this added up to loss of prestige, something much valued in academic life. Nonetheless, he accepted. If other men had made a downward move for the sake of their wife's academic career, we did not know of any. Decision made, Larry never expressed regret or complaint either at the time or during the rest of his life.

Fears of the academic periphery notwithstanding, in the years at Virginia Tech Larry not only continued to publish but produced some of his best known work in history and philosophy of science, consolidating his intellectual reputation. Among his articles and books were "A Confutation of Convergent Realism" (1981), *Science and Hypothesis* (1982) on how the history of philosophy of science had its own trajectory, quite different from the history of epistemology as generally understood, and *Science and Values* (1984) on adjudicating cognitive values. For conversation, he found the philosopher of statistics Deborah Mayo, Jarrett Leplin in the nearby 'research triangle' of North Carolina, as well as the visitors and seminars in the Science Studies Center invaluable in replacing his Pitt buddies.

The relevance of his work to broader issues was also becoming clear. With gusto, he entered the so-called 'science wars.' The legitimacy and authority of science which Larry so valued was being challenged by sociologists, literary theorists, and anthropologists, with potentially serious consequences. The relativism of this anti-science movement was unfounded, Larry argued. Science, although its theories could not be known to be true, and although scientists made many blunders, produced more reliable knowledge than any other intellectual endeavor.

Ironically at the same time, letters kept arriving from groups such as nurses hoping to learn how to achieve scientific status for their discipline by adopting one or another of the theories of scientific change then in play. Although Larry himself, in the wake of others such as Popper, Kuhn, and Lakatos, had proposed one of these theories of scientific change, he worried that the theories themselves did not begin to have evidential warrant, being based on a handful of case studies. This was the origin of what came to be known as the Virginia Tech project. Eight of us from Virginia Tech and elsewhere—both historians and philosophers—produced a catalog of the empirical claims made by the different theories of change. We then invited anyone knowledgeable, whether a scientist, a sociologist, a philosopher or a historian, to write a paper explaining why the case they knew best did or did not bear out one of those claims. The procedure was collegial, though historians of science bristled at the thought of being asked to test philosophical claims.

Along similar lines, when fellow philosopher of science, Michael Ruse, made the case against creationism using what Larry regarded as a bad argument that in the end

gives support to creationism, he leapt into the fray to explain why. It did not win him friends across the board, many believing creationism had to be defeated whatever the later fallout, but it was one of many times when Larry acted on what he believed.

I was busy too. I sent off my manuscript, *From Mineralogy to Geology: the Foundations of a Science*, to the University of Chicago Press. I was granted tenure. My next project—uncovering scientists' evolving ideas between 1750 and 1950 about why science evolved—was well under way, supported by the two major government funding agencies, the National Science Foundation and the National Endowment for the Humanities. Larry as always trusted me and supported my work in history of geology, technology, the history of the politics of science (and later food) even when he either did not understand its importance or was uninterested in the issues. Never did he try to divert me to topics that he thought more reasonable or more likely to advance my career.

We had a good social life. Living in the beautiful valley between the Blue Ridge and the Appalachians, was pleasant. To relax, Larry added local folk music festivals and playing the guitar, banjo, and dulcimer, to the strategy games he had played in Pittsburgh. I meantime was off hiking in the mountains. We could have stayed at Virginia Tech.

That was not to be however. Institution building is perilous. Creating the Science Studies Center had unleashed a micro-political storm. The Dean of Science and Humanities had diverted positions and money from existing departments (something I had already experienced to my dismay at Carnegie-Mellon). The new chair of History, who like the historians at Pittsburgh disdained history of science as marginal, fought to get them back. Larry's high salary, part of the practice just then beginning of hiring 'stars' to improve a department or university's rating, had not surprisingly generated resentment. We certainly made mistakes. So did others. Tempers flared, insinuations and rumors flew. Even normally steady and equable Larry was shaken. By the second half of the 1980s, the Dean had been forced to resign and Arthur Donovan, Larry and I had decided that the storm could not be ridden out. Once again, we went on the job market.

MOVING TO THE MIDDLE OF THE PACIFIC

In 1987, Larry accepted an offer to chair the University of Hawaii Philosophy Department. The department's strength, Asian philosophy, ranked even lower than continental philosophy in the opinion of analytic philosophers. With the mainland 2500 miles away across open ocean and email not even in its infancy, staying 'in the network' with colleagues looked difficult to impossible. If moving to Virginia Tech had been a demotion, this seemed as if we might be falling off the edge of the world. Friends chuckled at the thought of Larry under a swaying palm watching the waves rolling in, swopping philosophy for sunbathing.

Since Larry had forfeited much for me in our previous move, it was now my turn. As long as he was in charge, the chair of History told me, a position for an accompanying spouse would never be forthcoming. I started over in a part-time non-tenure track position in a marginal department (general science for future teachers) that the administration wanted to abolish as quickly as possible—an accompanying spouse dumping ground.

The Islands were exorbitantly expensive, the multi-racial population in which east Asian not Anglo culture predominated, thanks to the descendants of Hakka and Han Chinese, Japanese and Okinawans, Koreans, and Filipinos who had come as indentured laborers on the sugar and pineapple plantations, was not welcoming to mainlanders. I remember trudging up a hill on my daily walk, kicking one of the fist-sized invasive snails that are such a nuisance in the islands, reflecting guiltily that I was back to square one, and Larry too was worse off. We were going backward, not forward.

Once again, this initial angst proved to be not just overwrought but quite misplaced. We began to enjoy the intellectual life, not just the sun, sea, and sand so relentlessly promoted to tourists. Larry found philosophers of science, Ron Amundsen at the regional campus, Ron Pine at a community college good to talk to. The United States so over-produces Ph.Ds. that every university has some excellent faculty. Angus Graham, Roger Ames, and David Kalupahana in Chinese and Buddhist philosophy convinced Larry that these were intellectually serious endeavors misunderstood because of poor nineteenth century translations. Another colleague, an enthusiast for applied philosophy, encouraged Larry's growing interest in legal evidence. Then there were astronomers, vulcanologists, biologists, oceanologists and agronomists attracted by the physical environment. Many in the humanities specialized in east Asian Studies. For my part, I was fascinated by the historians, led by Jerry Bentley and distinguished visitors, who were figuring out how to write reputable world history in a place where courses limited to the 'West' made no sense.

Our careers did not seem to be suffering as we had anticipated either. Larry published his rebuttal to relativism in the form of a dialogue, *Science and Relativism* (1990). He was elected president of his professional society, the American Philosophical Association, Pacific Division.

I worked my way into a full-time, tenured, full professorship, and chaired the General Science Department as it morphed into a Biology Department. I was close to publishing my book on the history of ideas of scientific progress. As a hobby, I had begun exploring the many small eateries and mom n' pop stores locally, as well as the excellent and beautiful State Library, to understand the to-me exotic foods so important to the students and staff in the department. I came to understand that three diasporas, from the south Pacific, from Asia, and from Europe and the Americas, had each brought their food systems to these barren island and used them to survive and create a unique local culture.

After the predominant mainland use of food as a vehicle for snobbery, understanding food as a reflection of cultural history was refreshing, intriguing and challenging. I discovered a network of independent food scholars that stretched from England to the United States, from the Philippines to Mexico whose generous collegiality was a welcome change in academic reserve. My book, *The Food of Paradise: Exploring Hawaii's Culinary Heritage*, was about to be published by the University of Hawaii Press.

LEAVING ACADEMIA

We were, however, becoming increasingly restless in academia. Among our concerns about the direction it was taking were the disregard for undergraduate teaching; the rising cost of a university education for students and their families; the decline of collegiality that accompanied the star system; the inaccessibility of campus resources to the wider public; the increasing bureaucratization; and the uncontrolled proliferation of graduate programs and consequent growth of justifiably-disgruntled un- or under-employed Ph.Ds.

Our restiveness intensified during a visiting year in Princeton in the early 1990s. We had anticipated enjoying being back in the academic mainstream and the culture of the east coast. In many respects, our hopes were fulfilled. The Friday Seminars of the Davis Center of the History Department were a model of intellectual engagement, with their pre-circulated papers and unsparing but collegial debate. Meeting old and new colleagues visiting the School of Social Science in the Institute for Advanced Study for their year dedicated to science was a pleasure.

Other encounters were not so productive, however. The permanent members of the School of Social Science made it clear that that they had felt pressured to make science the topic for the year and that they were on the 'anti' side of the science wars. We visitors speculated about what was going on to create this dismal atmosphere. Larry was growing weary of rehearsing arguments against relativism. For my part, I was discouraged at the thought of publishing a book that showed how radically scientists' ideas about progress had changed over time. It could be too easily seen as yet more anti-science ammunition. I tried out on various faculty the alternative plan I was hatching of writing a world history of food using Hawaii as a model. They winced at the word 'food' and again at the word 'world.' Neither fitted their vision of good history.

And to our surprise, after Hawaii, the culture of the Princeton institutions now seemed tired and rather dull. Conversations all too often consisted of gossip about the elite universities of the coasts, western Europe and Israel or commentary on articles in the New York Times, the New York Review of Books, and the New Yorker. From the point of view of the Asia-Pacific region from which we now came, having this handful of publications circumscribe cultural life seemed a tad parochial.

We discussed our growing disenchantment with the seemingly unquestioned assumption that everything important academically happened in elite centers with the distinguished social scientist, Don T. Campbell, who frequently visited Princeton. He pointed out that among social scientists it was common knowledge that innovations were frequently introduced from the periphery, where there was less pressure to conform and more exposure to stimulating new influences. The problem was to get a hearing for the innovations in centers that condescended to the periphery. If that could be done, perhaps because of the reputation of the scholar on the margins, mutually beneficial interaction ensued. That, unfortunately, was not happening in the School of Social Sciences.

On our return to Hawaii, Larry and I mused that we had spent all our conscious years in educational institutions. Did we want to continue for another quarter of a century? Would we begin repeating ourselves? What new would we learn? Although Larry continued to hold as an ideal an integrated history and philosophy of science that informed the larger scientific enterprise, he saw no hope of that in the immediate future. He was not even sure he wanted to continue intellectual work. I was, by now yearning to have a go at a world history of food, laying bare the intellectual causes of major culinary changes. Larry had earned a good salary from a young age so that his university-matched savings plans were well funded. If we could earn a little money on the side, our savings would tide us over for the remaining years before we could claim the American safety nets of Social Security and Medicare. We were not tied to laboratories as so many scientists were. Was it time to leave academia, free up positions for younger scholars, and embark on a new adventure?

Yes, we decided. We assumed many late middle-aged academics like us with savings tucked away were considering something similar. It seemed not. A colleague from Harvard told me that he and some of his friends had also considered leaving academia in their 50s. But what about the loss of prestige and structure associated with a university position?

Never one for looking back once he had made a decision, Larry began planning our exit with me. He got advances on two popular books on the gap between the evidence and popular beliefs about risks, *The Book of Risks* (1995), and *Danger Ahead* (1995). Once again, he was happy to see his work have practical import when one was excerpted at length in *Good Housekeeping* in the same month that he gave his presidential address to the American Philosophical Association. I wrote seemingly endless encyclopedia articles, at least learning how to quickly summarize the essentials of a topic.

We had to find somewhere less expensive and less distant than Hawaii. We did not want to be nomads, nor to be confined to an expatriate community. We wanted an interesting location where we could become reasonably fluent in the language. The mainland United States? Hardly a new challenge. Europe? Too expensive. Asia? No way we could learn the language fast enough. Australia and New Zealand? Too far from family.

Mexico looked good. Prior visits to give talks at the National Autonomous University in Mexico City (the UNAM) had impressed us with the intellectual life and its pivotal position in the Spanish-speaking university world. The thought of seeing the Americas from the Spanish-speaking perspective was intriguing. We alerted colleagues at the UNAM to our decision, telling them that we would enjoy informal connections, such as participating in conferences or meeting graduate students, but that we were not interested in positions as we had decided to leave academia.

In 1996, we moved to Guanajuato, Mexico, a beautiful colonial town on the high central plateau, a place for several centuries at the center of the global silver trade, with architecture, libraries, and economic and political traditions that reflected this heritage. We set about learning Spanish, helped by doctoral students, particularly Godfrey Guillamin, who stayed with us exchanging lessons in return for conversation about his thesis.

Larry toyed with various ways to occupy himself. He wrote drafts of books on intellectual revolutions, the Copernican revolution, and political theory, but was not satisfied with any of them. He considered starting a business, an itch he had inherited from his entrepreneurial parents. An earlier stab at a computer game rental business had come to nothing, but something else might occur to him. Perhaps, given the risk books had brought lots of queries, consulting would be interesting and lucrative. Fiddling around with internet access with the computer specialist at the Guanajuato University Library was lots of fun. None of these ventures held his attention for very long.

What did hold his attention was comparative procedure in Mexican (Roman and Anglo-Saxon law). Helped by Mexican colleagues, particularly Juan Antonio Cruz Parcero and Edgar Aguilera, Larry, continuing his lifelong fascination with evidence and rational decisions, this time verdicts in criminal law, began working seriously on what he called legal epistemology. On our annual visit to Austin, Texas for libraries, shopping, and a base for visiting family, Larry introduced himself to the philosopher of law, Brian Leiter, at the University of Texas at Austin Law School, who not only introduced him to other philosophers of law but arranged for him to teach an accelerated course in jurisprudence in the Law School. Meantime, Ron Allen, Wigmore Professor of Jurisprudence at Northwestern University, who had once years before contacted Larry about a philosophical question, was back in touch with him. The two of them began on a series of jointly authored papers.

BACK TO ACADEMIA

So in 2001, Olbeth Hansberg, Director of the Philosophical Research Institute at the National University invited us to Mexico City and offered us both positions in the Institute. By this time, Larry was happy to accept. About that time, Spanish philosopher of law, Jordi Ferrer Beltran, visited the Instituto and was delighted

to discover in Larry a like-minded colleague. Before long Larry was organizing regular small conferences—all the better for intense conversational exchange—on legal epistemology. These brought together philosophers of law who might never have met otherwise from Mexico, Spain, France, Italy, Colombia, England, and the United States, including Amalia Amaya and Andrés Páez, the editors of this volume. Larry was particularly delighted that, unlike his work in philosophy of science that was largely ignored by scientists, lawyers took an interest in his work in legal epistemology.

Larry was now back from the periphery in two centers, the Institute which was a major hub in the network of Spanish-language universities, and the University of Texas at Austin Law School. This was something he had never anticipated on our move to Mexico. In 2008, about a decade after he had declared his departure from academia, Larry published *Truth, Error and Criminal Law* with Cambridge University Press.

I declined the offer from the Institute, since my research was just too far from its mission. I now felt self-confident enough to pursue intellectual work outside the academy. Meanwhile, while retaining my extra-university network of food scholars in the Anglophone world, I met Mexican historians and anthropologists of food in the seminars of the Anthropological Institute. Thanks to the generous Visiting Scholar status, which although unpaid, gave crucial library access first in Latin American Studies, then in History at the University of Texas at Austin, I was able to complete *Cuisine and Empire: Cooking in World History*. To my astonishment, food history had become trendy in history department micropolitics. Even geology was having its moment in history and philosophy of science circles.

During the Mexico years we made new friends, relished the excitement of working in new research areas, and had the enormous pleasure, thanks to our locations and invitations to Spain, Argentina, Panama, and Colombia, of seeing the Americas from a whole new perspective. It was a wonderful, invigorating fifteen years to top off our careers.

Then in 2012 Larry announced we were leaving Mexico, a unilateral decision that was completely out of character for him. Only later did I understand that he realized his Spanish, crucial for conversation, was slipping. Whether he understood it was the onset of dementia, something both his grandfather and father had suffered and he had always feared, I have no idea. Soon he resigned from UT Austin, though he continued to struggle to write. At her invitation, we moved closer to his daughter, and to his son too. Larry passed his remaining years largely peacefully and without ever knowing, or at least acknowledging, that anything was wrong.

CONCLUSION

Was Larry's unconventional career worth it? Not everyone would have found it so. The frequent moves, both permanent and for visiting appointments, took a toll in terms of time, money, and probably reputation. Remaining in one place, and hammering away at a single topic instead of switching research areas, might have meant more recognition of an official kind. Yet the journal Larry created and the departments in Pittsburgh and Virginia Tech that Larry established or helped establish still flourish. His works in philosophy of science and law are still cited. Legal epistemology is thriving in the Spanish-speaking world and the English-speaking one too. Larry always had time to puzzle over issues he believed to be important and to have at least some of them taken seriously in the extra-academic world. To have lived and worked (not just been a tourist or expatriate) in England, the mainland United States, Hawaii, Mexico, and (briefly) in Argentina and Spain was an honor. It was Larry's hope that his move from one university to another and from one country to others contributed to collaboration in the academic world, though that is for others to judge. Finally, Larry and I had changed in parallel ways over our fifty years together, we had made good friendships, we had remained equal partners, trusting each other, and making decisions jointly. The bet Larry and I had made on each other when we got together had paid off handsomely.

Quaestio facti. Revista Internacional sobre Razonamiento Probatorio
Quaestio facti. International Journal on Evidential Legal Reasoning
Sección: Homenaje a Larry Laudan
2026 I 10 pp. 25-28
Madrid, 2026
Marcial Pons Ediciones Jurídicas y Sociales
© Amalia Amaya
© Andrés Páez
ISSN: 2604-6202
Editado bajo licencia Reconocimiento 4.0 Internacional de Creative Commons

EL LEGADO DE LARRY LAUDAN Y EL FUTURO DE LA EPISTEMOLOGÍA JURÍDICA

Amalia Amaya
Universidad de Edimburgo
Andrés Páez
Universidad de los Andes

Como editores de esta sección especial de *Quaestio Facti* dedicada a la obra y legado del filósofo Larry Laudan, nos sentimos muy afortunados de haber conocido a Larry y de haberlo tenido como mentor, amigo y colega.

Larry fue un filósofo tremendamente generoso. Recordamos las largas tardes discutiendo con él, sin prisa, haciendo filosofía de una manera que contrasta de manera radical con el ritmo frenético de la academia actual. Siempre admiramos la manera en la que Larry siempre seguía el argumento hasta el final, así le llevara a posiciones que pudieran resultar comprometidas, por decir lo menos, o directamente impopulares.

Larry era un apasionado de los argumentos. Para él, los argumentos son el eje central de la práctica académica en todas sus dimensiones. En una ocasión, comentó que él recordaba argumentos que se habían presentado en congresos y seminarios hacía muchos años, pero muchas veces no se acordaba de quién los había presentado. Ahí estaban, los argumentos, listos para ser discutidos, examinados, suspendidos en el aire, como fórmulas matemáticas. Un artículo, decía, es «*one pointed argument*» y en las presentaciones en los congresos lo importante es el argumento, de lo que se trata es simplemente de presentar una idea que quieres discutir con los demás.

La relevancia de la comunidad para la práctica académica —y esto puede resultar quizá sorprendente— es, junto con la centralidad de los argumentos, otra de las

grandes enseñanzas de Larry. Organizó durante varios años mini-foros de epistemología jurídica, con pocas personas, los «*usual suspects*», como él los llamaba, lo que permitía conocerse bien y argumentar efectivamente. Para Larry la comunidad era un «test» importante para determinar qué ideas valía la pena seguir discutiendo y una brújula para orientar la propia investigación. Discutir, argumentar —¡de manera, a veces, infatigable!— eran para Larry la clave para un trabajo académico serio y placentero: permiten avanzar el conocimiento a la par que, sin duda para Larry, se pueden disfrutar enormemente.

El trabajo de Larry fue central para consolidar la epistemología del derecho como área de investigación y configurar una comunidad académica singularmente diversa. Su trabajo, nos parece, también fue pionero en apuntar a tres de las que son algunas de las direcciones centrales en las que sería importante avanzar en los próximos años.

(a) En primer lugar, para Larry era fundamental apoyar las tesis con datos empíricos obtenidos de diversas disciplinas. Sin duda, cuando Larry empezó a trabajar en filosofía del derecho, se usaban datos empíricos, sobre todo en estudios sobre el jurado, y empezaban a surgir trabajos sobre prueba que incorporaban resultados empíricos de la psicología cognitiva y social. Sin embargo, el uso por parte de Larry de estadísticas y resultados empíricos para evaluar tesis normativas y arreglos institucionales fue altamente novedoso. La insistencia en la relevancia de los datos empíricos para el quehacer filosófico es la tesis central de la filosofía basada en evidencia, de un naturalismo y pragmatismo que nos parece, abre nuevos horizontes que merece la pena explorar en el ámbito de la epistemología jurídica.

(b) En segundo lugar, Larry estaba firmemente convencido de que la filosofía de la ciencia tenía mucho que aportar a la filosofía del derecho, como su obra (y la de otros que han continuado construyendo puentes entre ambas disciplinas) pone de manifiesto. La epistemología de la prueba jurídica ha avanzado de manera importante gracias a la aplicación de métodos y teorías en filosofía de la ciencia. Sin embargo, consideramos que la epistemología jurídica es mucho más amplia que la epistemología de la prueba jurídica. Para empezar, la epistemología de la prueba jurídica se orienta hacia el pasado: trata de explicar hechos pasados, pero hay problemas epistémicos centrales también hacia el futuro, y no estamos hablando aquí de ciencia ficción, aunque algunos de los hechos que hemos vivido recientemente parecieran más propios de una película de este género. Como ha quedado vívidamente de manifiesto durante la pandemia, es importante poder validar el conocimiento de los riesgos, anticipar eventos y cursos de acción posibles, «predecir» las consecuencias de las distintas decisiones que se pueden tomar para estar en condiciones de enfrentarlas eficazmente. De manera más mundana, el derecho siempre ha estado preocupado por predecir: por ejemplo, por saber cuáles son las potenciales consecuencias de una política pública, de una nueva ley, o las probabilidades de que una persona convicta vuelva a delinquir si se le concede libertad condicional o de que un acusado escape a la justicia si se le otorga libertad bajo fianza. Las ciencias no solo tratan de explicar sino también de predecir. Nos parece que en la filosofía de la ciencia también pode-

mos encontrar herramientas importantes para avanzar, no solo en el conocimiento de hechos pasados, que es el objeto característico de la epistemología de la prueba jurídica, sino también en predecir los riesgos, las implicaciones y las consecuencias más probables de las decisiones tomadas en el presente.

(c) En tercer lugar, Larry logró algo realmente difícil: integrar comunidades académicas —la anglosajona y la iberoamericana— y propiciar un fructífero diálogo entre ellas. Recientemente, ha habido un interés creciente entre los epistemólogos en temas de epistemología jurídica, en una medida importantísima propiciado por el trabajo de Larry. En las últimas décadas, han surgido importantes colaboraciones entre filósofos de la ciencia y filosofía del derecho o juristas que trabajan en prueba y teóricos del derecho. Nos parece que sería necesario —en el espíritu de Larry— hacer un esfuerzo por integrar la filosofía del derecho y la epistemología y establecer puentes entre ambas comunidades.

(d) Finalmente, Larry en sus últimos trabajos dio un giro hacia la filosofía política —explorando por ejemplo la relevancia de nociones epistémicas para examinar el contrato social—. El giro hacia la epistemología política es posiblemente uno de los desarrollos más interesantes en epistemología y de mayor relevancia para la epistemología aplicada y, especialmente, para la epistemología jurídica. Nos parece que sería importante situar los temas de epistemología política en la agenda de investigación en epistemología jurídica.

Siguiendo los pasos de Larry, que nunca dejó de aprender cosas nuevas, quisiéramos explorar cuáles podrían ser las futuras líneas de investigación en epistemología jurídica. Hemos mencionado cómo el legado de Larry apunta a la necesidad de abordar cuestiones de epistemología experimental, conocimiento del futuro y epistemología política, así como de forjar lazos con la comunidad de epistemólogos y filósofos de la ciencia. Pero ¿en qué direcciones adicionales podríamos avanzar? Se nos ocurren tres frentes en los que la epistemología jurídica puede profundizar en discusiones que aún están en un estado incipiente.

En primer lugar, ha habido un énfasis casi exclusivo en las decisiones individuales y en los procesos cognitivos de jueces y testigos, pero el estudio de las decisiones grupales está subdesarrollado. Existen múltiples aportes de la epistemología y la psicología social, y del estudio de los procesos de investigación colectiva en la ciencia —donde la confianza y la mutua dependencia epistémica son fundamentales—, de los cuales se podría nutrir el derecho.

En segundo lugar, no podemos dejar de mencionar el aporte que puede hacer la epistemología aplicada al uso de la inteligencia artificial (IA) en el derecho. No solo se trata de hacer un control epistémico a los modelos de decisión automatizados que se usan en casos de libertad condicional, por ejemplo. La irrupción de la inteligencia artificial generativa (IAG) ha creado problemas dentro de las decisiones judiciales, no solo en términos de las sentencias inexistentes que se inventa la IAG y que aparecen en las decisiones de los jueces, sino en general, en la actitud desprevenida de

muchos agentes judiciales en el uso de grandes modelos de lenguaje (LLMs), que son usados como si fueran una fuente confiable de hechos. El uso de *deepfakes* de audio y video está destinado a crear una crisis sin precedentes en la valoración de la prueba, y la solución va a requerir una reevalución del peso y el riesgo epistémico de muchos tipos de evidencia.

En tercer lugar, parece necesario expandir la mirada de la epistemología jurídica, que se ha centrado de manera principal en un momento procesal (la etapa probatoria), en un actor jurídico (el juez o miembro del jurado), y en los sistemas jurídicos estatales, para embarcarse en el análisis de dimensiones epistemológicas del derecho relevantes en el contexto de otros roles jurídicos (como el de procurador, el legislador o el regulador); otras instituciones jurídicas distintas del proceso (como el sistema electoral) y en ámbitos supraestatales (tanto a nivel regional como internacional, por ejemplo, el estudio de la prueba en el contexto de las cortes del sistema internacional de justicia, problemas probatorios en situaciones de justicia transicional, o cuestiones epistemológicas relativas a la generación de conocimiento en contextos organizacionales internacionales, como la OMS o la OTAN). Las dimensiones epistemológicas del derecho abarcan una diversidad de cuestiones que van más allá de la epistemología de la prueba jurídica y que constituyen un área de investigación fecunda aún por explorar.

Queda mucho por hacer. Invitamos a los epistemólogos, filósofos políticos, de la ciencia y del derecho a continuar trabajando en la epistemología jurídica y continuar, de esta manera, la disciplina a la que el trabajo de Larry tanto contribuyó. En este número de la revista *Quaestio Facti*, los lectores encontrarán varias contribuciones que, tomando el trabajo de Larry Laudan como punto de partida, se ocupan de esa tarea.

Quaestio facti. Revista Internacional sobre Razonamiento Probatorio
Quaestio facti. International Journal on Evidential Legal Reasoning
Sección: Homenaje a Larry Laudan
2026 I 10 pp. 29-45
Madrid, 2025
DOI: 10.33115/udg_bib/qf.i9.23122
Marcial Pons Ediciones Jurídicas y Sociales
© Ernesto Saccardi*
ISSN: 2604-6202
Recibido: 28/02/2025 | Aceptado: 25/06/2025 | Publicado *online*: 19/11/2025
Editado bajo licencia Reconocimiento 4.0 Internacional de Creative Commons

LAUDAN'S META-RULE FOR EVALUATING EVIDENCE AND THE CASE OF EXPERT TESTIMONY

Ernesto Saccardi*
Department of Civilization and Forms of Knowledge,
University of Pisa, Italy.
e.saccardi@studenti.unipi.it

ABSTRACT: This article examines the challenges posed by expert testimony within the context of the US legal system, focusing primarily on the rules of evidence and judicial practices regulating admissibility—most notably Rule 702 of the Federal Rules of Evidence and the interpretive framework provided by the *Daubert* trilogy. Although many of the epistemological concerns discussed may have broader significance, the analysis is firmly grounded in the American adversarial tradition. Whitin this framework, this article aims to show that in the absence of effective criteria for establishing the reliability of expert testimony, the meta-rule for the admission of evidence proposed by Laudan in *Truth, Error and Criminal Trial* is unusable. Judges are laypeople burdened with the complicated task of distinguishing reliable experts from charlatans. However, from an epistemic perspective, this situation is extremely challenging. Indeed, due to the epistemic asymmetry between laypeople and experts, judges are in a very weak epistemic position to carry out their gatekeeping duty, and unfortunately the *Daubert* standard does not seem to be useful in this respect. What I would like to show is that using a meta-rule for evaluating evidence based on the concept of reliability without previously addressing the question of how to determine the reliability of expert testimony is putting the cart before the horse.

KEYWORDS: expert testimony; Larry Laudan; expertise; Daubert standard; social epistemology.

* I'm very grateful to Prof. Roberto Gronda for all his valuable suggestions. I am also grateful to the anonymous reviewers of the journal for their insightful comments on the initial draft of this article.

SUMMARY: 1. LAUDAN ON EVIDENCE.— 2. THE EPISTEMIC PROBLEMS OF THE GATEKEEPING FUNCTION OF JUDGES.— 3. THE INEFFECTIVENESS OF *DAUBERT* FACTORS.— 4. LAUDAN'S META-RULE AND THE EXPERT TESTIMONY.— 5. PROXIES OF EXPERTISE.— BIBLIOGRAPHY.

1. LAUDAN ON EVIDENCE

According to Laudan (2006), the criminal trial is an "epistemic engine", namely an essential tool by which societies try to find out the truth about a confused situation. The legal fact-finding process can be, indeed, embedded among the truth-seeking practices examined by epistemology. Nevertheless, the legal context can be an inhospitable ground for epistemologists due to the strict boundary between non-epistemic and epistemic values and *Truth, Error and Criminal Trial* is, indeed, conceived as a thought experiment. The central question of the book is a counterfactual one: «What would a criminal trial look like if its fundamental and overriding concern were to find out the truth about a crime?" (Laudan, 2006, p.123).

In the second part of the book, Laudan examines a series of rules and procedures that seem to interfere with the task of finding out the truth and avoiding error. If we want to find out the truth, we must condemn only those who are truly guilty and acquit only those who are truly innocent. Obviously, achieving 100% correct verdicts is a chimerical goal due to the fact that verdicts are products of human judgment and, therefore, are subjected to an unavoidable margin of error. Despite this, what epistemology can do is to try to identify some epistemically non-suspect rules and procedures that can help legal fact-finders in their complicated task of distinguishing the truly guilty from the truly innocent. In Laudan's words:

> Anything we can do to make it more likely that the truly guilty are seen to be guilty and that the truly innocent appear to be innocent reduces the overall likelihood of erroneous verdicts. It should be the principal function of the rules of evidence and procedure to ensure that a defendant's apparent guilt is as good an indicator of his true guilt or innocence as we can make it. (2006, p.119)

Laudan suggests an epistemic reframing of some of the rules of evidence and procedures effective in the US legal system. As Laudan correctly states, if we want the verdict at the end of a criminal trial to be as correct as possible, it is absolutely essential that juries have access to evidence of satisfactory quality and quantity. According to him, however, several of the current exclusions of evidence are not based on truth-seeking purposes but, on the contrary, «highly germane evidence—evidence that would indisputably reduce the overall rate of erroneous verdicts—is excluded for reasons having nothing to do with the search for the truth" (Laudan, 2006, p.120).

Laudan's extensive examination of numerous rules of evidence leads him to conclude that these rules are not founded on any concerns regarding truth. According to him, a significant number of rules of evidence are informed by the concern that jurors may erroneously evaluate evidence, failing to assign them the appropriate degree of

consideration. Indeed, many current rules of evidence are the consequence of the fact that courts are worried that jurors might evaluate the evidence wrongly. However, he claims, this fear is based on unrefined considerations of juries' psychology, and he sees such considerations as an obstacle to a truth-driven system of judicial adjudication. Consequently, these rules impede the effective functioning of the criminal trial, seen as an epistemic engine designed to establish an accurate reconstruction of the facts.

Laudan argues that to maximize the chances of factfinders reaching an accurate verdict, it is essential to establish truth-oriented admission rules for evidence. Accordingly, the question, as posited by Laudan, is how to construct a set of evidentiary rules that would enhance the probability of reaching an accurate verdict. To achieve this end, Laudan claims that the question of admissibility should be considered through the lens of two key concepts: *relevance* and *reliability*.

Laudan defines relevance in a straightforward manner, stating that a piece of evidence is relevant if, if true, it increases or decreases the likelihood of guilt of the accused. Reliability is, instead, a more complex matter, but, for the sake of my argument, I think we can simply define it as something we can rely upon in order to obtain a belief that is likely to be true. Finally, we can enunciate what Laudan (2006) calls «a simple rule for guiding our decision about the admission of evidence» (p. 121):

> *Laudan's Meta-Rule for Evaluating Evidence*: The triers of fact– whether jurors or judges in a bench trial– should see all (and only) the reliable, nonredundant evidence that is relevant to the events associated with the alleged crime. (Laudan, 2006, p. 121)

To which he adds the following "precautionary" *codicil* a few lines later:

> *Codicil*: Where there is compelling evidence that a certain type of relevant and reliable evidence is likely to be misconstrued by jurors and where there is likewise evidence that such misconstruals are not readily susceptible of correction by judge's instructions and the arguments of opposing counsel, such evidence should be excluded. (Laudan, 2006, p. 122)

Laudan shows that many admission rules for evidence are not informed by this *meta-rule* and, thus, these seem to be an obstacle to a truth-driven system of judicial adjudication. It is interesting to note, however, that Laudan spends very few words on expert testimony, and I think it would be worthwhile to examine whether and how his meta-rule can be used for guiding our decision about the admission of expert witnesses. Nevertheless, it is first necessary to discuss some epistemic problems that arise with regard to expert testimony. The following section will address these issues in greater detail.

2. THE EPISTEMIC PROBLEMS OF THE GATEKEEPING FUNCTION OF JUDGES

It is well known that the use of expert witnesses in civil and criminal trials is becoming increasingly common. Indeed, the judicial system can greatly benefit from scientific advances in order to reach a greater number of factually correct verdicts.

However, this progress is also giving rise to an increasing number of problems concerning the admissibility of such evidence. As things stand, in the US legal system, judges have a gatekeeping role. According to the Federal Rule of Evidence, judges must determine whether expert testimony can be considered reliable and whether it can be submitted to the jurors. Since the beginning, I would like to stress that, as Martini (2015) and Brewer (1998) argue, judges are laypeople charged with the complex task of determining the reliability of an expert witness. It is evident that a certain tension exists with regard to the evaluation of the reliability of scientific evidence by a person who lacks specific scientific training. In this sense, the issue can be seen as an instance of the more general problem of what Selinger (2011) calls "the recognition problem". This problem can be resumed with the question «How can a novice identify experts in a given field of expertise?». The issue arises because laypeople, by definition, appear to lack the capacity to establish directly the trustworthiness of experts.

This presents a similar challenge to that faced by judges in determining the reliability of expert testimony. Judges are not omniscient, and, obviously, they lack knowledge in many domains of expertise. No one is an expert in every domain. Consequently, it is unclear how they can ascertain the trustworthiness of a particular expert testimony. As Brewer (1998) pointed out, the question is how a scientifically untrained judge could be epistemically competent to perform his gatekeeping task. Experts are necessary because they provide useful information about the case that cannot be obtained directly by judges and juries. Indeed, if they were able to arrive at the same information on their own, then there would be no need to involve experts.

Furthermore, it is important to note that the legal context seems to be strongly characterized by the presence of unreliable experts. As many authors have pointed out (see, for instance, Huber, 1993 or Posner, 1999), one of the most problematic aspects of adversarial systems is the strong presence of partisan experts. The fact that many testimonies are given under compensation seems to create an additional difficulty for those who have to establish the trustworthiness of experts. Indeed, if it is already complicated for laypeople to recognize experts to trust, this becomes even more difficult in a context where the risk of encountering junk science is high. For all these reasons, it is not difficult to envisage scenarios in which the decision regarding the admission of expert witnesses is made on the basis of non-intellectual factors. In such cases, if the decision turns out to be correct, it appears possible to conclude that the accuracy of the outcome is just a matter of luck.

Consider the following example:

> *Case A*: Judge Smith must decide whether to admit the testimony of Dr. Johnson, a forensic physician who conducted research to determine the cause of death of a murder victim. Everything seems to be in order, and the conclusions the expert has reached seem to Judge Smith obtained through reliable methods. However, Judge Smith is a layperson who knows absolutely nothing about forensic medicine, and his assessment is based exclusively on what he is able to understand. His judgment is not an assessment of substance but is based solely on appearances. Judge Smith

> decides to admit Dr. Johnson's testimony. Fortunately, Dr. Johnson is actually a reliable witness, and his findings are indeed scientifically acceptable[1].

But there seems to be a problem here: is his decision based on well-founded reasons or is it just a case of epistemic luck[2]? In other words, is it just a matter of luck that the judge has a true belief?

Consider an alternative scenario:

> *Case B*: Judge Smith must decide whether to admit the testimony of Dr. Franklin, a forensic physician who conducted research to determine the cause of death of a murder victim. Everything seems to be in order, and the conclusions reached by the expert seem to Judge Smith obtained through reliable methods. However, as in case A, Judge Smith is a layperson who knows absolutely nothing about forensic medicine and his assessment is based exclusively on what he is able to understand. His is not an assessment of substance but is based solely on appearances. This time, unlike Dr. Johnson, Dr. Franklin is not a reliable witness, and he is just making the case for the side that is paying him through an argument that might appear "scientific" to people who aren't experts in his field.

Are we sure that Judge Smith will be able to distinguish a case of real science from one of junk science? Are we sure that a judge—who i recall is a layperson—will be able to distinguish a trustworthy expert from an untrustworthy one?

Judge Smith, as a layperson, is incapable of assessing directly the content of the testimony, and then he does not appear to be in the epistemic position to distinguish between a trustworthy expert and one who *only appears* trustworthy.

It could be noted that some of the described issues, such as the possibility of bias or rhetorical manipulation on the part of the expert, are more relevant to the fact-finder's evaluation of the testimony's *content* than to its *admissibility*. As we will see in the next paragraph, under Daubert, judges' gatekeeping function concerns the

---

[1] It is interesting to note the strong similarities between this case and the famous fake-barn problem discussed by Goldman (1976): "Henry is driving in the country-side with his son. For the boy's edification Henry identifies various objects on the landscape as they come into view. "That's a cow", "That's a tractor", "That's a silo", "That's a barn", etc. Henry has no doubt about the identity of these object; in particular he has no doubt that the last-mentioned object is a barn, which indeed it is. Each of the identified objects has features characteristic of its type. Moreover, each object is fully in view, Henry has excellent eyesight, and he has enough time to look at them reason- ably carefully, since there is little traffic to distract him. Given this information, would we say that Henry knows that the object is a barn? Most of us would have little hesitation in saying this, so long as we were not in a certain philosophical frame of mind. Contrast our inclination here with the inclination we would have if we were given some additional information. Suppose we are told that, unknown to Henry, the district he has just entered is full of papier-mach6 facsimiles of barns. These facsimiles look from the road exactly like barns, but are really just facades, without back walls or interiors, quite incapable of being used as barns. They are so cleverly constructed that travellers invariably mistake them for barns. Having just entered the district, Henry has not encountered any facsimiles; the object he sees is a genuine barn. But if the object on that site were a facsimile, Henry would mistake it for a barn. Given this new information, we would be strongly inclined to withdraw the claim that Henry knows the object is a barn. How is this change in our assessment to be explained?" (p. 772-773).

[2] For a clear and comprehensive explanation of the concept of "epistemic luck" see Pritchard (2005).

scientific validity of the methods employed, not the purity of the expert's motives. Issues such as bias or exaggeration are usually addressed during cross-examination. However, while the formal distinction between admissibility and the evaluation of the testimony's content may be clear in legal doctrine, it may be blurred in practice. The reliability of an expert's methodology cannot be evaluated in isolation from the content of the conclusions drawn. As Haack (2016) noted, in the 1997 General Electric Co. v. Joiner case, the second in the *Daubert trilogy*, the Supreme Court refined its earlier stance by affirming that a judge may exclude expert testimony when there is an excessive discrepancy between the underlying data and the conclusions drawn from it. In doing so, the Court eroded the distinction between admissibility and evaluation, pushing judges towards implicit judgements about the content of expert testimony rather than its methodology alone.[3] The epistemic problems concerning judge's gatekeeping task become even more complex when we consider the high level of specialisation that characterises the sciences. To illustrate, one might consider the field of medical sciences. A neurologist, for instance, addresses very different clinical situations than those handled by a neurosurgeon, and their respective domains of primary expertise are obviously distinct. It is easy to envisage scenarios in which a novice may be unable to identify instances of *epistemic trespassing*. According to Ballantyne (2019), we can define epistemic trespassers as «thinkers who have competence or expertise in make good judgment in one field, but move to other field where they lack competence» (p. 367). Therefore, returning to the legal context, from an epistemological point of view, a judge should not only assess the reliability of an expert but also whether his testimony falls within his domain of expertise.

But this type of problem does not only concern the gatekeeping task assigned to judges. A further complication seems to arise if we consider the issue of *peer disagreement*, which is a characteristic feature of adversarial systems. This issue is particularly relevant to the fact-finding task of juries, who must decide which expert to trust most when the parties present conflicting expert testimonies. The jury, in this case, faces the challenge of determining which of the two experts' testimonies is more reliable, given that the two experts will be providing contradictory evidence. This scenario appears to be an illustration of what Goldman (2001) designates as the *novice/2-expert problem*. In short, the question is whether a layperson can reasonably decide which expert is more trustworthy with respect to a particular issue. In the adversarial tradition, faith is placed in the idea that the competition between two competing testimonies will optimise the search for truth. However, as Goldman (1999) noted, this line of argument seems extremely problematic. Firstly, it operates under the assumption that the premises employed by the disputants are all true, a supposition that is dubious when considering the context of scientific testimony. Secondly, it presupposes that juries are able to correctly evaluate the relationship between the thesis proposed and the evidential support offered, which seems extremely difficult if we consider that jurors are laypeople.

[3] I would like to thank you an anonymous reviewer for encouraging me to clarify this point.

As previously stated, the fundamental epistemological problem underlying all these considerations is that laypeople usually do not have the epistemic competence to adjudicate who is an expert in a specific field and who is not. Judges are laypeople appointed to determine the reliability of expert testimony. Therefore, from an epistemological point of view, a problem seems to arise with regard to their gatekeeping function: How can a person who doesn't have any specific competence in a certain domain establish if someone is a trustworthy expert in that domain?[4]

3. THE INEFFECTIVENESS OF *DAUBERT* FACTORS

The first decision in the US legal system history about the involvement of expert witnesses in courts was in *Freye v. United States* (1923). In this case, the court decided to reject the results of a then-new systolic blood pressure deception test on the grounds that novel scientific evidence did not have a sufficient degree of general acceptance in the relevant scientific community to be admitted into evidence. This decision established that expert testimony must be generally accepted in the relevant scientific community in order to be admissible. As Martini (2015) pointed out, this standard "seems to be grounded on the idea that good science tends to generate consensus around established facts and methods" (p. 4). This appears to be intuitively correct. Ideally, a hypothesis is presented to a scientific community, and through a complex (and non-linear) process of trial and error, is slowly accepted. However, Martini (2015) emphasises, «the problem with the use of consensus criteria in law is that too often consensus is only a *byproduct*, not a cause of good, and therefore, court admissible, science» (p. 4). Indeed, it is not possible to exclude the possibility that scientific community consensus is formed around spurious scientific ideas as a result of extra-scientific and non-neutral factors. Moreover, even if consensus were a good criterion for assessing the reliability of scientific expert testimony, its formation process would still be too slow to accommodate its purpose.

More than fifty years after Freye v. United States, in 1975, the common-law evidentiary rules were for the first time codified in the Federal Rules of Evidence (FRE). Rule 702, which pertains to scientific knowledge, stipulates that a qualified expert witness may be admitted provided that his testimony is relevant to the subject matter of the trial. However, as Haack (2014) correctly observes, the absence of any reference to the "general acceptance" mentioned in Freye v. United States leaves the question of whether FRE 702 had replaced Freye unresolved.

[4] As noted by Tuzet (2023), alongside these epistemic considerations, other semiotic ones can be added. The experts' language is a technical one and it contains a rich amount of signs that are often impossible to understand to laypeople. This semiotic problem generates a justification problem «fact-finders cannot form a justified belief about the relevant matter, nor justify the acceptance of an expert testimony, insofar as they do not understand it» (p. 1)

One of the most important decisions concerning the admissibility of expert testimony was in the 1993 case *Daubert v. Merrell Dow Pharmaceuticals, Inc*., the first of the three United States Supreme Court cases (so-called *Daubert trilogy*) that articulated the current rule of evidence regarding the admissibility of expert witness testimony. One of the central aspects of Daubert is that it is emphasised that in order to admit expert testimony, the judge must not only assess its relevance but also its reliability. As I have already said, under Daubert, the trial judge is seen as a gatekeeper who must guarantee the reliability of experts presented before the jury. To fulfil his gatekeeping function, the judge must assess whether a given scientific testimony is the product of reliable principles and methods. In the sentence, indeed, Justice Blackmun argued that in order to evaluate the reliability of an expert, a judge must look at the methodology used and not at the conclusion he reached. In order to accomplish this task, the court listed several "factors" that might be considered by the judge to determine whether the expert's methodology is admissible:

1) Whether the theory or technique has been tested or is capable of being tested;

2) Whether the theory or technique has undergone peer review and has been published;

3) The known or potential error rate associated with the specific technique;

4) The existence and maintenance of standards for the control of its operation;

5) Whether the theory or technique is widely accepted by the relevant scientific community.

In *General Electric & co. v. Joiner* (1997*)*, the second of the trials composing the *Daubert* trilogy, the court clarified the previous ruling by stating that expert testimony could be declared inadmissible if the judge finds the 'analytical gap' between the data and the conclusions too large.

Eventually, in *Kumho Tire Co. v. Carmichael* (1999), the court extended *Daubert*'s gatekeeping function beyond scientific evidence to encompass all expert testimony. In December 2000, the *Daubert* trilogy was finally codified in FRE 702.

The *Daubert* standard has been the subject of extensive discussion and has been strongly criticised in several respects[5]. Haack (2014), for instance, showed how, from the first ruling composing the *Daubert* Trilogy, there is a marked use of the term "scientific" in its honorific sense. According to her, the term "scientific" and its cognates have, indeed, an honorific use that generates an illegitimate identification between what is scientific and what is reliable. This idea becomes clear when we consider that the court assumes the existence of a unique "scientific method," understood as the only reliable method of inquiry, used by all scientists and by only scientists. But this seems to be extremely problematic for at least two reasons. Firstly, there is not

[5] For a detailed analysis of the history and key issues surrounding the admission of expert testimony in court, see, among others, Hilbert (2019).

a unique mode of rational inference of the kind envisaged by the court. Secondly, it is important to remember that the reliability of the method does not guarantee that the method will produce true results; a "reliable method," in its technical sense, is a method that generates consistent results. These considerations lead to the conclusion that the image of the scientific enterprise on which the *Daubert* Trilogy rulings are based is extremely different from the actual scientific practice.

Further critiques focus on the broader institutional and socio-political implications of the *Daubert* ruling. Edmond and Mercer (2004) argue that the Daubert rulings have created an 'exclusionary ethos' that disproportionately disadvantages plaintiffs, particularly in toxic tort litigation, by making it harder to admit expert evidence. According to their analysis, this shift reflects a convergence of judicial and corporate interests in promoting restrictive evidentiary standards. Similarly, Mnookin (2010) notes that courts frequently fail to apply *Daubert* criteria rigorously in cases involving traditional forensic sciences, revealing a double standard in judicial practice. Other commentators have highlighted the paradoxical outcomes of applying *Daubert*. Giocoli (2020), for example, points out that, in certain antitrust cases, the testimony of Nobel Prize-winning economists has been rejected under *Daubert*. Finally, it is important to note that some scholars have recognised that the *Daubert* standard represents a significative effort in attempting to address some fundamental issues associated with expert testimony. Bernstein (2008), for instance, claims that the *Daubert* framework shifts the focus of judicial attention towards the epistemic basis of testimony. This provides a partial solution to the problem of testimony being distorted by partisan experts, although this criterion alone is not enough to solve the problem completely.

Nevertheless, the focus of this article pertains to the epistemic issues discussed in the previous paragraph. As Martini (2015) pointed out, «Both the *Frye* test and rule 702 […] shift the problem from evaluating the validity of expertise to evaluating the validity of the evidence presented» (p. 6). I fully agree with him when he says that the *Daubert* standard completely ignores the problem of how a layperson can identify trustworthy experts. Although the *Daubert* standard formally restricts the judge's gatekeeping role to assessing methodological soundness, in practice, this often entails engaging with the content of the expert's conclusions. However, a judge, as a layperson, doesn't possess the necessary competence to evaluate the evidence presented by an expert directly; we rely on experts exactly because we are unable to do it alone. The *Daubert* standard, focusing mainly on the method and technique employed by the expert, assumes that judges are able to make a direct assessment of the content of the testimony, and, in doing so, it fails to recognize the epistemic asymmetry between experts and laypeople.

In order to carry out their gatekeeping task, the judges should establish whether an expert is a *reliable source of information*. Nevertheless, if they, as laypeople, cannot do this through a process of direct evaluation of the evidence presented, the only alternative seems to be to address the epistemic problem of how laypeople can recog-

nise experts. In the next section, I will argue that in order to be consistent with the Laudan's idea that the trial should aim primarily at the truth, we need to establish an effective way to establish the reliability of expert testimonies without ignoring the epistemic asymmetry between experts and laypeople.

4. LAUDAN'S META-RULE AND THE EXPERT TESTIMONY

Finally, we can return to Laudan's meta-rule and its difficulties in dealing with expert testimony. The core idea of Laudan's legal epistemology is that the primary aim of a judicial system should be to reach a judicial truth that is as close as possible to the factual one. Laudan believes that epistemology can exert an analytic effort useful to identify all those norms that obstruct this process of truth discovery. As mentioned in § 1, Laudan in *Truth, Error and Criminal Trial* (2006) dedicates many pages to examining some exclusionary rules present in the US judicial system. According to his analysis, many of these are not based on truth-seeking purposes but, on the contrary, represent an obstacle to the process of discernment between the truly guilty and the truly innocent. As already stated, in order to raise the likelihood of a correct verdict, Laudan proposes the *meta-rule* mentioned in the first paragraph.

In my opinion, Laudan's idea that trials should aim at finding out the truth about a confused situation is correct. Accurate fact-finding it's not just an intellectual aspiration. On the contrary, it is a prerequisite for achieving justice. Although procedural fairness and institutional legitimacy are indisputably pivotal to the effective functioning of a legal system, they run the risk of becoming mere formalities if they are detached from the aim of arriving at truthful conclusions. Although in practice it is unrealistic to always arrive at factually correct verdicts, truth must continue to serve as a regulative ideal. However, although Laudan grasps this important point, difficulties emerge when contemplating the employment of his meta-rule for the regulation of expert testimony. In this section I would like to stress that in the absence of epistemically reliable criteria for distinguishing between expert witnesses who can be trusted and those who cannot, Laudan's meta-rule seems to encounter difficulties when used to guide the decisions about the admissibility of expert testimony.

The crux of the issue lies in the fact that the meta-rule proposed by Laudan is founded on two key concepts, namely "relevance" and "reliability"; jurors should see all and only reliable and relevant evidence. Unfortunately, in the case of expert testimony, the inability of judges, as laypeople, to recognise who is a trustworthy expert witness poses a significant challenge in the practical application of this rule. Indeed, the problem is that in order to apply this meta-rule to the case of expert testimony, it is necessary that the question of how to establish the reliability of experts has already been resolved. The situation could be partly improved if there were effective criteria by which judges could evaluate the trustworthiness of expert testimony. In the present state of affairs, as previously outlined, the judge performs a gatekeeping function, determining the admissibility of expert testimony prior to its presentation

to the jury. In this evaluative role, the judge is required to ground their decision on FRE 702. However, as previously discussed, the *Daubert* standard encounters significant difficulties and is completely ineffective in addressing the epistemic challenges identified in § 2. In the absence of sound epistemological criteria for determining the reliability of expert testimony, Laudan's meta-rule appears to be of limited practical application in such cases. Indeed, in order to admit only relevant and reliable evidence, it is necessary to establish efficient methods for evaluating the reliability of expert testimony. In other words, it is essential to identify effective methods of recognising trustworthy experts before this meta-rule can be fully utilised.

It's important to note that immediately after stating this meta-rule, Laudan (2006) suggests that "there is no epistemic reason not to leave reliability decisions entirely in the hands of the jury itself" (p. 121). Hence, Laudan argues, it's possible to delete the term "reliable" from the meta-rule presented above. According to him, this would facilitate the application of his meta-rule in concrete judicial practice. Nevertheless, I don't think that this consideration affect the outcome. Even if we remove the term "reliable" from the meta-rule, a decision about the reliability of experts should still be made, and, at this point, it would be entirely in the hands of the jurors. But jurors are laypeople just like judges, and from an epistemological point of view, the problem remains unresolved: how can a person with no technical competence distinguish a trustworthy expert from a charlatan? Removing the term "reliable" from the meta-rule does not resolve the issue; it merely postpones it.

Personally, I believe that the task of determining whether an expert is trustworthy should remain in the hands of the judges. If the decision were left to the jury, they would be in the difficult position of having to evaluate and make their judgment based on multiple conflicting expert opinions, with no minimal assurance of the reliability of any of the testimony. In this scenario, the decision on the reliability of an expert should instead be made at the height of the cross-examination phase, which further complicates matters. On the one hand, there is the problem of peer disagreement, as mentioned in the first paragraph. On the other hand, if it is already difficult to come up with criteria that can be used by judges, it is even more difficult to do so for juries. However, the most problematic aspect remains that, given that juries are composed of laypeople, the selection of trustworthy testimony could plausibly be based on non-epistemic factors, making, for example, an expert's rhetorical skills more relevant than his actual expertise.

In any case, what is important to note is that beyond these counterfactual scenarios, the current situation is highly problematic. Judges perform their gatekeeping role on the basis of epistemically ineffective criteria, and there is no guarantee that judges will ground their judgement about the reliability of experts on well-founded reasons. Despite this, I believe that by leaving the burden of determining the reliability of expert testimony to them, it is possible to identify epistemically less suspect criteria that can help judges in their complicated task of identifying reliable experts.

As already stated, in my opinion, Laudan's idea that the trials should aim at finding out the truth about a confused situation is correct. However, in order to maximize the chances of success of this enterprise, it is necessary to attempt to remove what obstructs this process of truth discovery. The use of scientific experts in trials is clearly an indispensable element in order to maximise the chances of obtaining an accurate verdict. Unfortunately, the current criteria by which the reliability of expert testimony should be assessed do not seem to be effective. Given the importance of identifying effective strategies to guide judges' decisions on the admissibility of expert testimony, some progress needs to be made. But what solutions can we hope for? The problems associated with the gatekeeping role of judges are many and varied, and it is likely that only systematic and collaborative work across different disciplines can improve the current situation. Nevertheless, for the purpose of this paper, given the centrality of the underlying epistemological problems that I have highlighted in § 2, I believe that it would be useful to give a brief overview to one of the most interesting solutions to the problem of expert recognition that social epistemology has developed in the latest years.

5. PROXIES OF EXPERTISE

The problem of recognition of experts is obviously an extremely interesting issue for epistemologists. This problem concerns all cases in which we base our judgements, our beliefs, and our actions on what the experts tell us. As Cody (1992) pointed out, in modern societies we pervasively rely on others' testimony. However, this does not seem to be perfectly clear-cut, and some tensions seem to emerge with respect to the epistemic asymmetry concerning the relationship between experts and non-experts. Hence, the problem of how can laypeople identify trustworthiness experts is central to many practices in our society and, as already stated, from an epistemological point of view, the case of the recognition of expert witness testimony by judges is just an instance of a more general problem. Consequently, I believe that it may be useful to consider whether the epistemological debate surrounding the recognition of experts may in some way be beneficial in order to identify new tools by which to assess the admissibility of expert testimony in court.

Over the last 20 years, many scholars have proposed what we can call the "proxies strategy". The core idea of this proposal is that we can identify a series of proxies (or "indicators") through which laypeople can assess the reliability of experts. Obviously these indicators must be available to laypeople and must not presuppose the ability to make a merit assessment.

The idea of using the proxies strategy in the debate on the role of judges as gatekeepers was put forward by Martini (2015). Since then, however, there seems to have been a lack of follow-up research. As we will see shortly, this is likely due to the fact that the application of this strategy in the legal context poses challenges that are not easily resolvable. However, given the centrality of the epistemological issues in

relation to the admission of expert witnesses in court, I believe that looking at the answers that social epistemology has provided to the problem of the recognition of experts can be useful in improving our current rules for the admission of expert testimony.

Following Grundmann (2025), we can distinguish between *positive* proxies and *negative* proxies. The latter are indicators that may lead one to suspect that the expert testimony may not be reliable. The former, on the contrary, can be taken as indicators of the "good quality" of a particular expert testimony.

Anderson (2011) sustains that we can distinguish between two kinds of negative factors: those that focus on the honesty of a particular testimony and those that focus on the epistemic irresponsibility.

According to her, factors that may indicate that an expert is not being honest include: *conflicts of interests*; *previous episodes of scientific dishonesty* (e.g., plagiarism, faking data, etc.); *claiming misleading information* (e.g., cherry picking data, the misleading use of statistics in order to manipulate the interpretation of results, etc.); or *misrepresentation of scientific opponents*.

To assess the epistemic irresponsibility of experts, Anderson proposes instead indicators such as *evasion of peer-review*; *dialogic irrationality* (e.g. continuing to support theories after they have been refuted and without responding to the objections); and *the promotion of crackpot theories*.[6]

Obviously, these criteria are not immune to criticism. Grundman (2025), for instance, points out that the absence of negative indicators does not guarantee that a putative expert is an actual expert. Consequently, to improve the likelihood that a proxies list effectively evaluates an expert's reliability, we need to supplement it with positive indicators.

There is no definitive and uncontroversial lists, but we can consider the following proxies to be among the most commonly cited:[7]

a) *Track Record*;

b) *Dialectical Competence*;

c) *Reputation*;

d) *Internal Consistency*;

e) *Discrimination Ability*;

f) *Pertinence*.

Let's take a brief look at each of these indicators. One factor that can indicate the reliability of a putative expert is his *track record* of predictions. Theoretically speaking, even a layman, who by definition has no special knowledge in a particular field,

[6] For a more detailed analysis see Anderson (2011, p.147-148).

[7] For alternative lists of indicators see Shanteau *et al.* (2002), Martini (2020) or Grundman (2025).

can judge whether an expert's predictions have come true. However, as Grundman (2025) pointed out, not all disciplines are predictive, and many experts do not make predictions. At the same time, it is far from uncontroversial to affirm that laymen are capable of judging whether a particular prediction has been confirmed.

Goldman (2001) defines *dialectical competence* as the ability of experts to deal with objections and questions. *Reputation* is instead made up of two distinct elements: experts' general credentials (e.g., having a Ph.D., being affiliated with prestigious institutions, being a university professor) and the consensus among peers around their theses. As mentioned above, consensus is a problematic factor. On the one hand, the fact that a particular theory enjoys a certain degree of acclaim in a particular community of experts does not mean that the theory is valid. On the other hand, in many disciplines, consensus is not easily achieved due to the presence of a contentious methodological status (consider, for instance, the case of psychology or economics).

Another positive indicator of expertise is the *internal consistency* of putative experts' judgments. Indeed, it would be completely anomalous, and we would be inclined to distrust someone who claims to be an expert if he or she made inconsistent judgements.

Discriminatory ability coincides instead with the expert's ability to discriminate between similar but not equivalent cases (see Hammond, 1996).

Finally, *pertinence* coincides with the expert's attitude towards making judgments within his field of competence (see Walton, 1989 or Reiss, 2008). However, as noted above, it is not easy for laypeople to recognise the cases of epistemic trespass, i.e. when an expert makes a judgement outside her area of expertise.

All these indicators, as we have seen, appears problematic when considered in isolation. Indeed, those who have spoken in favour of this solution stress the importance of assessing the reliability of an expert by using all the proxies. In this way, if a given expert satisfies all or almost all of them, he is likely to be a reliable expert. Unfortunately, the potential use of these proxies in the debate on the admissibility of expert testimony becomes more complex when we consider the specificity of the legal context. For example, the negative indicator of conflict of interest is problematic when we consider that all experts are hired by the parties and, therefore, have an interest in confirming one version rather than another. This depends on the substantial differences that exist between science and judicial practice. As Haack (2014) emphasised, an advocate's business, unlike the business of a genuine scientific inquirer, «is to make the strongest possible case that this—his side's—answer is the true one; so he will be most effective if he selects and emphasizes whatever evidence favours the proposition in question, and ignores or plays down the rest» (p. 33). But this kind of problem does not seem to characterise only the negative indicators: let's consider, for instance, the case of dialectical competence. In a context such as the legal one, a lawyer will have every interest in choosing an expert with good rhetorical skills. This proxy, which in ordinary contexts can be useful in the difficult task of distinguishing genuine experts from charlatans, can become a false clue when considering the par-

ticular type of skills and abilities by which experts are selected in the legal context. To these problems others of pratical nature are added: even if the strategy of proxies proves to be effective, how can we use it in concrete? Although some progress has been made, the literature about the practical application of the proxies strategy is still green. Guerrero (2016) argued in favour of the possibility to provide what he calls an "incentive alignment strategy", i.e. a strategy that aims to align the interests of experts and non-experts trough a reward-penalty system in order to provide reliable performance of experts. He concludes that it would be possible to design what he calls a "Database of Experts", i.e. a database in which all licensed experts are registered according to a classification criterion that follows the normal separation between areas of expertise. Theoretically speaking, it would be possible to include in this database not only all the information regarding qualifications, credentials and academic affiliations, but also whether the theses proposed by an expert are well regarded within the relevant expert community, as well as evidence of his epistemic integrity. In this way it would be possible for a layperson to evaluate an expert also on the basis of his veracity and epistemic integrity.

Obviously some issue arises. Firstly, as Guerrero himself admits, it remains unclear how we can establish the quality of expert performances without relying on others experts' judgment. Secondly, it is quite complicated to understand how an idea like this can become a reality. Normal social practices, with which applied epistemology is involved, are composed by more complex situations compared to those that we find in ideal contexts with which philosophy usually works.

The problem of expert recognition is certainly far from being solved. The proxies strategy is certainly one of the most interesting proposals that has been put forward in the last years. However, much work remains to be done. As we have seen, the problem becomes even more complex when we consider the particular context in which the problem of expert testimony is immersed. The adversarial system has very specific rules and values, and this makes even more arduous to resolve the epistemic issues surrounding the gatekeeping role of judges. The application to the legal context of the solutions proposed by epistemology to the problem of expert recognition seems to be indeed conspicuously hampered by the values and normal practices that are involved in our judicial systems.

Unfortunately, as I have argued above, the persistence of these epistemic questions about the gatekeeper role of judges significantly hampers the practical utilization of Laudan's meta-rule for evaluating evidence. Nevertheless, I believe that, in line with the central thesis of Laudan's legal epistemology, we can see how epistemology can produce an analytical effort in order to trace and diagnose some epistemological problems concerning the gatekeeping function of judges. As things stand, there is still a gap between the extant literature concerning the epistemic problem of recognising experts and that addressing the admissibility of expert witnesses in court. A definitive solution to the problem of expert testimony will probably never be found. However, an improvement of the actual situation is only possible through a joint ef-

fort of different disciplines, approaching the problem from different angles. Indeed, a productive way to apply the conceptual resources of epistemology would be to use them not only to identify the issues that shape certain social practices (including the challenge of expert testimony) but also to suggest potential solutions for improving the current state of affairs. This mutual interaction between epistemology and social practice would be of considerable benefit to both. The latter would gain significantly from the analytical attitude and conceptual tools that have long been central to philosophy. At the same time, when faced with everyday problems, epistemology—and philosophy more broadly—can carve out a new role in public discourse, providing both critical insights and normative guidance.

BIBLIOGRAPHY

Anderson, E. (2011). Democracy Public Policy and Lay Assesment of Scientific Testimony. *Episteme 8(2)*, 144-164. https://doi.org/10.3366/epi.2011.0013

Ballantyne, N. (2019). Epistemic Trespassing. *Mind, 128(510)*, 367-395. https://doi.org/10.1093/mind/fzx042

Bernstein, D. E. (2008). Expert Witnesses, Adversarial Bias, and the (Partial) Failure of the Daubert Revolution. *Iowa Law Review, 93(451)*, 101-137. https://ssrn.com/abstract=963461

Brewer, S. (1998). Scientific Expert Testimony and Intellectual due Process . *The Yale Law Journal*, 1535-1681.

Coady, C. A. (1992). *Testimony: a Philosophical Study.* Clarendon Press.

Edmond, G. and Mercer, D. (2004). Daubert and the Exclusionary Ethos: The Convergence of Corporate and Judicial Attitudes towards the Admissibility of Expert Evidence in Tort Litigation. *Law & Policy, 26(2)*, 231-257. https://doi.org/10.1111/j.0265-8240.2004.00011.x

Giocoli, N. (2020). Why do US judges reject antitrust experts? In P. C. Esposito, *Economics in legal reasoning* (pp. 101–117). Palgrave-Macmillan.

Goldman, A. I. (1976). Discrimination and Perceptual Knowledge. *Journal of Philosophy, 73(20)*, 771-791. https://doi.org/10.2307/2025679

Goldman, A. I. (1999). *Knowledge in a social world.* Oxford University Press.

Goldman, A. I. (2001). Experts: Which ones should you trust? *Philosophy and Phenomenological, 63(1)*, 85–110. https://doi.org/10.2307/3071090

Grundmann, T. (2025). Experts: Who are they and how can laypeople identify them? In M. STEIN, *Oxford Handbook of Social Epistemology.* Oxford University Press.

Guerrero, A. (2016). Living with ignorance in a world of experts. In *In Perspectives on Ignorance from Moral and Social Philosophy* (pp. 156-185). Taylor and Francis.

Haack, S. (2014). *Evidence matters: Science, Proof, and Truth in the Law.* Cambridge University Press.

Haack, S. (2016). Mind the Analytic Gap!: Tracing a Fault Line in Daubert. *Wayne Law Review, 31(3)*, 16-42. https://ssrn.com/abstract=2839027

Hammond, K. R. (1996). *Human Judgment and Social Policy.* Oxford University Press.

Hilbert, J. (2019). The Disappointing History of Science in the Courtroom: Freye, Daubert and the Ongoing Crisis of "Junk Science" of Criminal Trials. *Oklahoma Law Review, 71(3)*, 759-821. https://digitalcommons.law.ou.edu/olr/vol71/iss3/3

Huber, P. W. (1993). *Galileo's revenge: Junk Science in the Courtroom.* Basic Books.

Laudan, L. (2006). *Truth, error, and criminal law: An Essay in Legal Epistemology.* Cambridge University Press.

Martini, C. (2015). The Paradox Of Proof And Scientific Expertise,. *Humana.Mente Journal of Philosophical Studies, 28*, 1-16. https://www.humanamente.eu/PDF/Issue28_Papers_Martini.pdf

Martini, C. (2020). The Epistemology of Expertise. In M. F. Henderson, *The Routledge Handbook of Social Epistemology* (pp. 115-122). Routledge.
Mnookin, J. (2010, 75(4)). The Court, the Nas and the Future of Forensic Science. *Brooklyn Law Review*, 1209-1275. https://brooklynworks.brooklaw.edu/blr/vol75/iss4/10
Posner, R. (1999). The law and economics of the economic expert witness. *Journal of Economic, 13*, 91–99. https://www.jstor.org/stable/2647119
Pritchard, D. (2005). *Epistemic Luck.* Oxford University Press.
Reiss, J. (2008). *Error in Economics: Toward a More Evidence-based methodology.* Routledge.
Selinger, E. (2011). *Expertise: Philosophical reflections.* Automatic Press.
Shanteu, J. e. (2002). Performance Based assessment of expertise: how to decide if someone is an expert or not. *European Journal of Operation Research, 136(2)*, 253-263. https://doi.org/10.1016/S0377-2217(01)00113-8
Tuzet, G. (2023). Expert signs and legal burdens. *International Journal for the Semiotics of Law, 36(1)*, 159–183. https://dx.doi.org/10.1007/s11196-022-09932-3
Walton, N. (1986). Reasoned use of expertise in argumentation. *Argumentation, 3*, 59–73 https://doi.org/10.1007/BF00116417

Cases cited

Daubert v. Merrell Dow Pharmaceuticals 509 U.S. 579 (1993).
Frye v. United States 293 F. 1013, D.C. Circ., (1923).
General Electric Co. v. Joiner, 522 U.S. 136 (1997).
Kumho Tire Co. v. Carmichael, 526 U.S. 137 (1999).
United States of America. *Federal Rules of Evidence. https://www.uscourts.gov/forms-rules/current-rules-practice-procedure/federal-rules-evidence*

Quaestio facti. Revista Internacional sobre Razonamiento Probatorio
Quaestio facti. International Journal on Evidential Legal Reasoning
Sección: Homenaje a Larry Laudan
2026 I 10 pp. 47-65
Madrid, 2025
DOI: 10.33115/udg_bib/qf.i10.23123
Marcial Pons Ediciones Jurídicas y Sociales
© Andrés Páez
ISSN: 2604-6202
Recibido: 01/03/2025 | Aceptado: 10/06/2025 | Publicado *online*: 19/11/2025
Editado bajo licencia Reconocimiento 4.0 Internacional de Creative Commons

LA NEGLIGENCIA INSTITUCIONAL FRENTE AL RIESGO EPISTÉMICO DE LOS SESGOS COGNITIVOS EN EL DERECHO

Andrés Páez
Universidad de los Andes
apaez@uniandes.edu.co

RESUMEN: El riesgo epistémico en el derecho es la posibilidad de cometer errores o aceptar conclusiones falsas debido a la incertidumbre fáctica y a los valores no epistémicos que rodean la toma de decisiones. Uno de los factores que contribuye a la distorsión de las decisiones en todos los momentos procesales es la influencia innegable de los sesgos cognitivos. En este artículo explico cómo estos contribuyen a aumentar el riesgo epistémico y exploro las posibles razones por las que los sistemas judiciales no han tomado medidas significativas para eliminar esta fuente de error. Al final analizo el problema de cómo atribuir responsabilidades a los diferentes agentes que integran el sistema judicial ante una actitud que solo puede ser catalogada como negligencia epistémica.

PALABRAS CLAVE: Riesgo epistémico; Sesgos cognitivos; Responsabilidad colectiva; Responsabilidad institucional; Negligencia epistémica.

INSTITUTIONAL NEGLIGENCE REGARDING THE EPISTEMIC RISK OF COGNITIVE BIAS IN LAW

ABSTRACT: Epistemic risk in law is the possibility of making mistakes or accepting false conclusions due to factual uncertainty and non-epistemic values surrounding decision-making. One of the factors that contributes to the distortion of decisions at all procedural stages is the undeniable influence of cognitive biases. In this article, I explain how these biases increase epistemic risk and explore

possible reasons why judicial systems have not taken significant measures to eliminate this source of error. Finally, I analyze the problem of attributing responsibility to the different agents within the judicial system in the face of an attitude that can only be described as epistemic negligence.

KEYWORDS: epistemic risk, cognitive biases, collective responsibility, institutional responsibility, epistemic negligence.

SUMARIO: 1. INTRODUCCIÓN. - 2. ¿Qué es el riesgo epistémico?. 3. ¿En qué sentido aumentan los sesgos cognitivos el riesgo epistémico?. 4. ¿Por qué se tolera ese tipo de riesgo en el derecho? 5. ¿Es posible atribuir responsabilidades por los daños generados por ese riesgo? 6. Conclusiones. REFERENCIAS.

1. INTRODUCCIÓN

El riesgo epistémico es la posibilidad de cometer un error o de aceptar una conclusión falsa debido a la incertidumbre fáctica y la intrusión de valores no epistémicos en los procesos de inferencia o decisión basados en evidencia. Como disciplina empírica, el derecho está expuesto al riesgo epistémico en diversos momentos de los procesos judiciales. Larry Laudan fue pionero en la discusión del riesgo epistémico en el derecho, analizando sus posibles causas y las intervenciones y cambios que ayudarían a mitigarlo. En su libro, *Truth, error, and criminal law*, Laudan (2006) presenta de manera muy clara los factores procesales e institucionales que determinan la aparición y la distribución de errores en el establecimiento de los hechos. En sus propias palabras, su intención es «proponer cambios en las reglas existentes para eliminar o modificar aquellas que resulten ser obstáculos significativos para llegar a la verdad» (p. 3). Laudan muestra, además, que uno de los elementos fundamentales en el riesgo de error es la injerencia de valores no epistémicos en el diseño de los procesos probatorios. La baja tolerancia de una sociedad a las falsas condenas, por ejemplo, es un elemento que trasciende la relación entre evidencia e hipótesis y que tiene consecuencias para la probabilidad de que se cometan errores en los procesos.

Las fuentes de error en el derecho son muchas y muy variadas. En este ensayo me ocuparé de los riesgos epistémicos que surgen de los sesgos cognitivos que, como lo demuestra la literatura reciente, afectan negativamente los procesos de razonamiento y decisión de jueces, jurados y tribunales. Me interesa no solamente identificar los momentos procesales en los que estos pueden interferir en el razonamiento probatorio, sino también entender por qué los sistemas judiciales no han implementado medidas efectivas para controlar esta fuente de errores a pesar de que hoy en día son ampliamente conocidos. Saber de un riesgo y no tomar medidas para prevenir sus causas es una actitud negligente cuyos orígenes me interesa explorar. Igualmente, quiero utilizar el riesgo epistémico generado por los sesgos cognitivos para explorar cómo se podrían atribuir responsabilidades individuales cuando existe un caso de negligencia epistémica institucional.

El artículo está dividido en cuatro secciones principales. En la siguiente sección explicaré en más detalle en qué consiste el riesgo epistémico y cómo la comprensión filosófica actual de ese concepto es mucho más compleja que el simple riesgo inductivo definido en términos estadísticos. En la sección 3 explico por qué los sesgos cognitivos aumentan el riesgo epistémico a la luz de esta forma ampliada de entender el concepto. En la cuarta sección examino varias causas posibles de la negligencia institucional ante el claro riesgo epistémico generado por los sesgos cognitivos, y en la sección final presento una forma en que se podría responsabilizar a los diferentes actores de la rama judicial por su actitud negligente frente a este problema.

2. ¿QUÉ ES EL RIESGO EPISTÉMICO?

En su sentido original, el riesgo epistémico es equivalente al riesgo inductivo (Hempel, 1965). En términos muy generales, la idea de riesgo involucra tanto la probabilidad como las consecuencias de un evento indeseable. En cuanto a la probabilidad, el riesgo está asociado al *valor de la expectativa estadística* de que ocurra un evento indeseable de cierto tipo (Rasmussen, 1975). En cuanto a las consecuencias, el riesgo dependerá de cómo se juzgue la *severidad* del evento. En muchos casos, será necesaria una evaluación social o individual de la gravedad de las consecuencias del evento a partir de las actitudes y expectativas de las personas involucradas. El riesgo adquiere una dimensión epistémica cuando el evento indeseable es la decisión de aceptar una proposición falsa a través de un proceso de inferencia inductiva (Collins, 1996). Cuando el proceso de inferencia es de naturaleza estadística, el resultado epistémicamente indeseable es obtener falsos positivos y falsos negativos a partir de la evidencia, los llamados errores tipo I y II, respectivamente.

En el derecho, los falsos positivos y los falsos negativos corresponden a las falsas condenas y las falsas absoluciones, respectivamente. El valor de la expectativa estadística de que ocurra cualquiera de estos dos errores está determinado por el estándar de prueba adoptado en un sistema jurídico, y por factores como la distribución real de inocentes y culpables entre la población de imputados y el uso correcto de la evidencia disponible. Como lo señala Laudan (2006), la meta de todos los sistemas judiciales del mundo es reducir la ocurrencia de este tipo de errores, pero al mismo tiempo es necesario reconocer que estos son inevitables. Bajo el supuesto evidente de que los errores de tipo I (falsas condenas) son más graves y costosos que los de tipo II (falsas absoluciones), se han creado múltiples mecanismos doctrinales y prácticos para minimizar los primeros a costa de los segundos. Encontrar la razón o proporción adecuada entre ambos tipos de errores es un problema que trasciende el ámbito puramente probatorio y que depende de los valores y principios sociopolíticos de una sociedad; es decir, es un problema que tiene una dimensión pragmática inevitable. No entraré a discutir aquí el problema de la distribución de errores, al cual Laudan le dedica una parte importante de su obra, porque en este artículo me interesa introducir una forma más amplia y actual de entender el concepto de riesgo epistémico.

Tras la publicación de «Inductive risk and values in science», de Heather Douglas (2000), el concepto de riesgo epistémico se ha complejizado y se ha extendido mucho más allá del riesgo inductivo y la generación de errores tipo I y II. En su concepción original, el riesgo epistémico aparece en el paso final de un razonamiento, cuando se hace la transición inferencial de la evidencia a la aceptación o el rechazo de la hipótesis en consideración. Douglas identifica elementos valorativos adicionales en el *proceso* de investigación empírica, los cuales contribuyen a que surjan resultados diferentes a partir de la misma evidencia. En particular, ella señala que el diseño metodológico y la recopilación, clasificación e interpretación de los datos implican decisiones que están cargadas de valores no epistémicos. Douglas resume el punto de la siguiente manera:

> (1) ¿Es la metodología utilizada lo suficientemente sensible para detectar el fenómeno de interés o es demasiado sensible a las disrupciones y el ruido? (2) ¿Los datos están siendo caracterizados discriminando suficientemente entre el rango de posibles resultados o están siendo clasificados de manera demasiado fina? (3) ¿Es la muestra lo suficientemente grande para detectar el resultado o es tan grande que el fenómeno de interés se va a perder en el *dataset* (particularmente si el *dataset* no es lo suficientemente representativo)? Estas no son estrictamente hablando relaciones inversas entre falsos positivos y falsos negativos, sino que requieren del juicio del científico en su toma de decisiones: si arriesgar el error en una dirección o en otra (2017, p. 85)[1].

Así, dos científicos pueden llegar a conclusiones diferentes frente a los mismos hechos dependiendo del nivel de riesgo de error que cada uno esté dispuesto a aceptar durante el proceso de investigación, y de la metodología, los conceptos y las variables que adopte (Wilholt, 2009). La discrepancia no depende de cómo decidan distribuir los errores al final del proceso de acuerdo con un estándar de prueba, sino de desacuerdos legítimos sobre aspectos que pertenecen a la práctica científica misma. Estos desacuerdos no son epistémicos porque ambas opciones han sido validadas por la comunidad científica como procedimientos aceptables en la búsqueda de la verdad, sino que surgen por diferencias en la conceptualización e interpretación de los fenómenos y en la forma en que las diferentes metodologías detectan sus propiedades. Además, como lo señalan Biddle y Kukla (2017), muchas veces las discrepancias no ocurren por decisiones individuales, sino que los diferentes valores y estándares están embebidos en las instituciones dentro de las cuales trabajan los científicos.

Los elementos que señala Douglas aparecen claramente en los procesos judiciales que recurren a pruebas periciales (Coloma *et al.*, 2024). Por ejemplo, cuando se utiliza una prueba estadística para establecer una relación de causalidad general, esta no se rige exclusivamente por criterios epistémicos, ni siquiera por prácticas comúnmente aceptadas dentro de la comunidad científica, sino que también interviene el nivel de riesgo epistémico aceptable para cada perito, como advierte Douglas. El desacuerdo ocurre dentro del seno mismo de la práctica científica y no es posible decidir cuál perito está utilizando una práctica más adecuada porque no hay una que sea la práctica «correcta».

---

[1] Todas las traducciones de textos publicados originalmente en inglés son propias.

En términos generales, el daño principal en esta versión ampliada del riesgo epistémico sigue siendo el mismo que en el riesgo inductivo tradicional, a saber, la aceptación de creencias falsas, pero ahora generalizado a todas las etapas del proceso de investigación. Podemos llamar a este resultado indeseable un *daño alético*. Los daños aléticos pueden tener diversas causas. Por una parte, hay causas valorativamente neutrales, como hacer una inferencia lógica incorrecta o excluir variables cruciales debido a que se desconocen aspectos importantes del fenómeno objeto de investigación. En estos casos no hay una decisión metodológica o procedimental de por medio. Por otra parte, la práctica científica sí exige tomar decisiones no solo metodológicas sino también referentes a la conceptualización, producción e interpretación de la evidencia. Estas decisiones sí están cargadas de valores no epistémicos, como veremos a continuación [2].

Una primera decisión que se debe tomar en cualquier investigación empírica es cómo conceptualizar un problema de investigación y como operacionalizar los conceptos empíricos y teóricos resultantes. En las ciencias sociales y en algunas ciencias aledañas, operacionalizar un concepto es relacionarlo con posibles observaciones, además de saber cómo incluirlo dentro de un constructo teórico. Por ejemplo, la operacionalización del término «carcinogénico» en un estudio epidemiológico afecta no solo las conclusiones a las que llegue la investigación, sino también el diseño experimental y el análisis teórico de los resultados empíricos. Una definición muy amplia o muy restringida del concepto puede llevar a conclusiones totalmente diferentes. Este tipo de decisiones tienen repercusiones en el derecho. Si las conclusiones de los estudios epidemiológicos son defendidas por dos peritos enfrentados en un proceso jurídico, se creará una situación en la que no será posible zanjar el asunto exclusivamente a partir de la evidencia disponible. Ambas estarán sustentadas en prácticas científicas aceptables. En el derecho también ha habido discusiones en torno a la manera en que se deben conceptualizar los hechos del caso. Por ejemplo, la forma en que muchos códigos jurídicos definen la agresión sexual y la violación incluye elementos de fuerza y violencia. Esto ignora el hecho de que muchos casos de agresión sexual ocurren en el hogar y que el violador es alguien que la víctima conoce bien. La ausencia de violencia a menudo se utiliza para desestimar las acusaciones de violación (Tuerkheimer, 2017; Páez y Matida, 2023). Una conceptualización más amplia del concepto tendría efectos significativos sobre la manera en que se establecen los hechos en estos casos. La discrepancia conceptual producto de una diferencia valorativa puede aumentar el riesgo de error al permitir que una violación sin signos de violencia quede impune.

Otra decisión importante en el proceso de investigación es la de incluir o excluir variables que se encuentren en el límite de una medición, Por ejemplo, Biddle (2007) muestra cómo en un estudio del antiinflamatorio Vioxx, que fue retirado del mercado debido a que aumentaba el riesgo de enfermedad cardiovascular, se utilizó un

[2] La discusión que sigue está basada en gran parte en Biddle y Kukla (2017).

estándar de prueba muy alto para decidir si una muerte podía ser clasificada como de causa conocida, lo cual disminuyó el número de casos que sirvieron de evidencia contra el uso de Vioxx. Por otra parte, en muchas ocasiones los investigadores deciden favorecer ciertas variables desde el comienzo y excluir otras. En los procesos de modelado causal, por ejemplo, muchas veces es necesario excluir variables para que el problema sea manejable. Esto se hace de manera *ad hoc*, con base en supuestos y a partir del conocimiento general sobre el caso, pero sin que haya un principio metodológico universalmente aceptado para hacer la selección de variables. Ambas situaciones ocurren también en el derecho, donde es necesario llevar a cabo un proceso de depuración y selección del material probatorio antes de presentar el caso ante un jurado o un juez. Por ejemplo, si se tiene tanto video como prueba testimonial de un incidente ocurrido hace varios meses, un juez puede optar por utilizar solo el primero, dada la falibilidad de la memoria a largo plazo; sin embargo, otro juez puede considerar que la prueba testimonial aporta elementos contextuales que un video no puede proporcionar.

Finalmente, diferentes investigadores pueden tener diferente tolerancia al *ruido* en las decisiones (Khaneman *et al.*, 2021). El ruido es la alta variabilidad o dispersión en una serie de decisiones. Ante casos muy similares, las decisiones también deberían ser muy similares. La variabilidad o dispersión ocurre cuando hay decisiones muy diversas frente a casos parecidos. La variabilidad es impredecible porque no obedece a un factor determinado. Además, es posible conocer cuál es la dispersión de una serie de decisiones sin saber cuál es la decisión correcta. El ruido se diferencia del sesgo en que este sí es una desviación predecible. Por ejemplo, un juez que tome decisiones muy diferentes en casos muy similares está tomando decisiones ruidosas, mientras que uno que siempre juzgue a favor de las personas de raza blanca está sesgado. El ruido no solamente se puede atribuir a individuos. Una institución puede ser más o menos tolerante al ruido en las decisiones de sus integrantes. La importancia del ruido en las decisiones judiciales salió a la luz principalmente a raíz de la publicación de *Criminal sentences: Law without order*, de Marvin Frankel (1973). Desde ese entonces se han documentado numerosos casos de decisiones ruidosas por parte de los jueces, y una indiferencia generalizada frente al problema por parte del aparato judicial[3]. Como veremos más adelante, algo similar ocurre con el efecto nocivo de los sesgos cognitivos.

[3] Entre los muchos estudios que se hicieron en los Estados Unidos en los años posteriores a la publicación del libro de Frankel se encuentran: Partridge y Eldridge (1974), Austin y Williams (1977) y Clancy *et al.* (1981). Aunque en los años ochenta se introdujeron algunos lineamientos para disminuir las disparidades en las penas impuestas, estos fueron abandonados a comienzos de este siglo.

3. ¿EN QUÉ SENTIDO AUMENTAN LOS SESGOS COGNITIVOS EL RIESGO EPISTÉMICO?

El concepto de sesgo cognitivo fue introducido por los psicólogos israelíes Kahneman y Tversky a comienzos de los años setenta del siglo pasado[4]. Un sesgo cognitivo es una desviación sistemática, involuntaria e inconsciente de una norma o de un estándar de racionalidad al emitir un juicio perceptual o conceptual, al recordar un evento o al hacer una predicción. No se trata de un error fortuito, sino de comportamientos que ocurren consistentemente bajo circunstancias similares, y que por lo tanto son predecibles y replicables. Son causados principalmente por el uso de heurísticas, atajos que utilizamos en el procesamiento de información, pero también son producto de las limitaciones de nuestra capacidad cerebral de procesar información, de influencias emocionales, morales y sociales, y de distorsiones en el proceso de almacenamiento y búsqueda de información en la memoria (Páez, 2021). No deben ser confundidos con los llamados «sesgos sociales», como el racismo, el sexismo, la aporafobia, etc. Estos no surgen de nuestra constitución como seres racionales, sino que tienen un origen social y cultural, y son altamente sensibles al contexto, lo cual hace que sus efectos sean altamente variables y poco predecibles (Brownstein, 2018; Brownstein y Saul, 2016a, 2016b). En lo que sigue solo me ocuparé de los sesgos cognitivos, aunque es evidente que los sesgos sociales también son una fuente importante de error en la investigación científica y en los procesos judiciales. Sin embargo, el problema de los sesgos sociales requiere un tratamiento teórico diferente (Arcila-Valenzuela y Páez, 2024).

El estudio del efecto negativo de los sesgos cognitivos en el derecho cuenta ya con una larga tradición[5]. Desde la apertura de una investigación hasta el momento de la imposición de la pena existen múltiples ocasiones en que los sesgos de los diferentes agentes judiciales, y las heurísticas que utilizan, interfieren negativamente en sus razonamientos y decisiones[6]. Es evidente que sesgos como el sesgo de confirmación afectan la valoración de la evidencia y en consecuencia la aceptación o rechazo de la hipótesis del caso (Lidén, 2023). Dicho sesgo, que consiste en darle mayor credibilidad a la evidencia que concuerda con las creencias previas del sujeto, también afecta a los miembros de los jurados, en los países donde existe esta figura. Estos adoptan muy rápidamente una posición acerca de la culpabilidad o responsabilidad del imputado y valoran la evidencia a la luz de esa posición. En experimentos con jurados simulados, la evidencia muestra que el sesgo de confirmación lleva a una rápida polarización en sus posiciones (Myers y Lamm, 1976).

---

[4] Kahneman *et al.* (1982) ofrece un tratamiento comprehensivo del tema.

[5] *Véase*, por ejemplo, Saks y Kidd (1980) y Weinstein (2003).

[6] Las heurísticas también pueden tener efectos positivos, agilizando el razonamiento. La cuestión de cuándo son benéficas y cuándo no ha dado lugar a profundos desacuerdos en la literatura. Véase Gigerenzer (2006) para una defensa de su utilidad.

Otro sesgo que aumenta la probabilidad de error en el momento de la decisión sobre los hechos es el sesgo retrospectivo (Fischhoff *et al.*, 1978). Cuando se tiene conocimiento de un hecho, existe la tendencia a pensar que este era mucho más predecible de lo que en realidad era a partir de la información disponible. En términos más técnicos, es la tendencia a sobreestimar ex post las probabilidades ex ante. El resultado es que los jueces tienden a asignarle mayor responsabilidad de la debida a los imputados por negligencia penal (Rachlinski, 1998). Así, los sesgos cognitivos se convierten en un elemento que potencia el riesgo inductivo tradicional.

Pero el efecto de los sesgos ha sido detectado en muchos otros momentos procesales. Como vimos en la sección anterior, uno de los elementos que genera un riesgo epistémico en una investigación es la forma en que es conceptualizada una situación. La llamada heurística de representatividad (Kahneman y Tversky, 1973) puede potenciar el riesgo de error en ese momento procesal. En la vida diaria a menudo debemos decidir a cuál categoría pertenece un evento o una persona. Es muy frecuente que en el proceso de categorización nos dejemos guiar por alguna propiedad sobresaliente y superficial de los eventos o las personas, saltando rápidamente a la conclusión de que estos pertenecen a una categoría sin tener en cuenta evidencia que puede apuntar en la dirección opuesta. La heurística de representatividad contribuye fuertemente a que los jueces y jurados utilicen estereotipos inexactos al darle sentido a los hechos del caso (Bordalo *et al.*, 2016). A su vez, el uso de estereotipos refuerza el sesgo de confirmación porque los decisores van a darle más peso a las pruebas que refuercen el estereotipo[7].

Las decisiones metodológicas que aumentan o disminuyen el riesgo de error epistémico no solo ocurren en las técnicas y procedimientos de los peritos en los procesos. Dentro de la práctica jurídica misma existen formas diferentes de recaudar pruebas, de interrogar, de identificar testigos, entre muchos otros procedimientos, y en cada una de estas prácticas se deben tomar decisiones metodológicas. Dependiendo del tipo de metodología escogida, puede aumentar o disminuir la probabilidad de que los sesgos cognitivos contribuyan al riesgo de error. Por ejemplo, la creación de perfiles psicológicos de las personas que han cometido un delito es una práctica común en las investigaciones penales como apoyo en la identificación de sospechosos. Sin embargo, el momento de la investigación en el que los investigadores tienen acceso al perfil puede afectar negativamente sus razonamientos al reforzar estereotipos y activar el sesgo de confirmación (Marshall y Alison, 2007).

Adicionalmente, incluso en las causas valorativamente neutrales de los daños aléticos, como hacer una inferencia lógica o probabilística incorrecta, el riesgo epistémico es elevado por sesgos como el efecto ancla, la heurística de disponibilidad o igno-

[7] Los estereotipos se entienden como la atribución de un atributo positivo o negativo a una persona solamente por su pertenencia a un grupo social, profesional o cultural. El estereotipo puede ser justificado o no, en cuyo caso es más un prejuicio explícito.

rar la tasa base[8]. Aunque en estos casos no intervienen decisiones metodológicas que cambien el riesgo de error epistémico, sí hay decisiones institucionales que pueden contribuir a que se cometan este tipo de errores más frecuentemente. Por ejemplo, la manera en que se presenta la información estadística durante un juicio puede aumentar o disminuir el chance de que se activen sesgos conectados con nuestras capacidades de pensamiento probabilístico. También sería deseable que se implementara una mejor formación en lógica y estadística en las facultades de derecho y en las escuelas de formación judicial, pero eso tiende a ser más la excepción que la regla.

Existen docenas de sesgos cognitivos adicionales que han sido identificados en la literatura psicológica y es fácil ver cómo pueden tener efectos nocivos en diferentes momentos de los procesos judiciales. Sin embargo, no me interesa ofrecer aquí un tratamiento exhaustivo del tema, sino más bien mostrar que los efectos negativos de los sesgos cognitivos son lo suficientemente conocidos en el derecho como para plantear la pregunta acerca de las razones por las que son tolerados o ignorados desde un punto de vista institucional.

4. ¿POR QUÉ SE TOLERA ESE TIPO DE RIESGO EN EL DERECHO?

Hasta hace un tiempo, el estudio del efecto nocivo de los sesgos cognitivos en el derecho era un asunto de especialistas en ámbitos académicos. Sin embargo, durante los últimos años se ha vuelto más común encontrar discusiones sobre el tema en las facultades de derecho y en las escuelas de formación judicial. Hablar del sesgo de confirmación, por ejemplo, ya no suscita miradas de extrañeza. Estos avances en la educación de estudiantes, abogados y demás miembros de la rama judicial no han venido acompañados de un sentido de urgencia para introducir cambios en las normas procesales, en la estructura de las investigaciones y los procesos de prueba, y en la forma de tomar decisiones sobre los hechos. No ha ocurrido lo mismo en el caso de los sesgos sociales como el racismo o el sexismo, donde sí es fácil detectar esfuerzos por introducir reformas en el sistema jurídico, tal vez debido a las implicaciones políticas de dichos sesgos[9]. La pregunta que me interesa explorar es por qué no ha habido un esfuerzo equivalente en el caso de los sesgos cognitivos si el riesgo epistémico que generan está ampliamente documentado y el daño que producen es comparable al de los sesgos sociales. Aunque discutiré varias posibilidades teóricas a partir de la evidencia disponible, lo más probable es que el estado actual de la cuestión obedezca a una combinación de los factores que explico a continuación. En todo caso, sea cual sea la razón más poderosa, quiero mostrar que existe una negligencia institucional innegable frente al problema.

---

[8] Véase Páez (2021) para una explicación detallada de estos sesgos cognitivos asociados con la probabilidad.

[9] Gillis (2021) y O'Brien y Grosso (2020) resumen algunas de las reformas procesales que se han llevado a cabo para combatir el sexismo y el racismo en los procesos judiciales en los Estados Unidos, respectivamente.

Una primera posibilidad es que la ausencia de cambios se daba al deseo de preservar otros principios jurídicos que son considerados más valiosos y fundamentales. Se podría argumentar que es razonable no introducir cambios profundos que afecten principios como el *stare decisis*, la publicidad, la contradicción y —sobre todo en el derecho penal— el principio de inmediación. Sin embargo, también existen argumentos sólidos en contra de preservar estos principios a la luz de la evidencia psicológica. En el caso del *stare decisis*, Jois (2009) argumenta que preservarlo es preservar una historia de decisiones sesgadas, y es además una manera de reforzar sesgos como la preferencia por el *statu quo*, el peso excesivo que les damos a las opciones por defecto, la tendencia a seguir las decisiones de otros, y en general, la tendencia conservadora que reflejan muchos sesgos cognitivos. A menudo, el uso del *stare decisis* se ve afectado por la heurística de disponibilidad, y los casos más notorios reciben mayor atención que los más relevantes, así los primeros hayan sido cuestionados por la calidad de su razonamiento (Kuran y Sunstein, 1998). Finalmente, el *stare decisis* fomenta el conformismo con normas imperfectas, alimentado por el sesgo del statu quo (Hirsch, 2004).

Por su parte, el principio de inmediación ha sido criticado porque activa varios sesgos cognitivos conocidos. Por ejemplo, aumenta el sesgo de confirmación (Lidén, 2023) y el efecto halo (Nisbett y Wilson, 1977), que es la tendencia a juzgar positivamente el testimonio de un testigo a partir de su forma de hablar o de vestir. También se activa el sesgo de actualidad (*recency effect*), que hace que la información presentada al final de un juicio oral tenga más peso en la decisión (Miller y Campbell, 1959). Finalmente, el principio de inmediación presupone la habilidad, por parte de jueces y jurados, de detectar rasgos como la honestidad y la credibilidad a partir de los gestos y la forma de hablar de los testigos. Existe abundante evidencia de que los seres humanos carecemos de esa habilidad en interacciones limitadas con otras personas (Buller & Burgoon, 1996; Shieber, 2012; Wellborn, 1991). Además, generalmente en esas interacciones se activa el error de atribución fundamental, que es la tendencia a explicar el comportamiento de los demás basándose en factores internos como la personalidad, mientras que se subestiman las influencias situacionales externas que podrían estar afectando su comportamiento (Ross, 1977).

Una razón adicional para que no haya habido más cambios podría ser que muchas normas procesales protegen los derechos jurídicos de los imputados en los procesos penales y estas no admiten reformas fundamentales, incluso si son obstáculos en la búsqueda de la verdad. El problema, como Laudan lo señaló en varias ocasiones, es que algunas de esas normas les otorgan a los imputados protecciones que trascienden el acuerdo social sobre la repartición de errores representado por el estándar de prueba: «Muchas de las actuales reglas de evidencia son torpes esfuerzos por, o tienen el efecto de aumentar el beneficio de la duda que ya le otorga el estándar de prueba al imputado […] Se sigue que, una vez establecido el estándar de prueba, todas nuestras deliberaciones subsecuentes deben estar guiadas por la preocupación por reducir los errores y no por cambiar su distribución» (Laudan, 2006, pp. 122-123). El punto no es que haya un conflicto fundamental entre la protección de los derechos jurídicos de

los imputados y los cambios procesales que lleven a la mitigación de los sesgos en las normas, sino más bien que las reglas no deben interferir con la función del estándar de prueba.

Hay otras razones mucho menos defendibles para la inacción frente al riesgo epistémico de los sesgos cognitivos. En muchos casos, los sistemas tienen un apego injustificado por reglas procesales arcaicas que hace mucho tiempo han debido ser modificadas a la luz de la evidencia empírica que demuestra su inoperancia en el establecimiento de los hechos[10]. También se aduce en ocasiones que los jurados, en los países donde se usa esta figura, son incompetentes y se deben implementar reglas procesales para guiarlos en su tarea[11]. Estas reglas, sin embargo, pueden tener efectos perjudiciales porque facilitan la aparición de sesgos, poniendo obstáculos en la búsqueda de la verdad. Finalmente, otra explicación inadmisible para la ausencia de reformas significativas es la resistencia de los jueces a adoptar reglas que limiten su poder de decisión, así estas tengan el propósito de prevenir la aparición de sesgos. Como afirma Schauer,

> Debido a que la mayoría de las reglas de evidencia son esencialmente exclusionarias, y como los jueces (como el resto de nosotros) no quiere aislarse de información potencialmente útil, muchos jueces tratan al derecho probatorio como un estorbo contraproducente que debe ser evitado cuando sea posible (2006, p. 1090).

Es necesario admitir, sin embargo, que las cortes no sienten ninguna presión para reformar y mejorar la calidad de sus procesos de decisión porque no reciben retroalimentación alguna acerca de la calidad de sus razonamientos (Posner, 1999). Por el momento queda como una pregunta abierta si sería mejor tener jueces que conserven su poder de decisión pero que estén mejor entrenados para evitar los efectos de los sesgos cognitivos, o si sería preferible tener reglas diseñadas para evitar que estos aparezcan y que limiten su libertad. Mi posición siempre se ha inclinado hacia la segunda posibilidad, pero en el fondo es una pregunta empírica.

Existen otras razones para la inacción que surgen de lo que llamaré la *excusa institucional*. Como lo señala Miranda Fricker refiriéndose a los sesgos sociales, «el poder del individuo para mejorar la situación es muy limitado. Los tipos de cambios procesales que creemos que pueden remediar la situación e impedir que se repita son fundamentalmente cambios *institucionales*» (2016, p. 48). Dichos cambios institucionales, argumenta Fricker, tienen que ser impulsados por individuos virtuosos que

---

[10] Las «reglas procesales arcaicas» cambian de un país a otro, pero los litigantes saben muy bien cuáles son. En mis conversaciones con Laudan, él solía mencionar ejemplos de todas las latitudes. Un ejemplo ilustrativo es la regla ancestral en el Reino Unido que les prohíbe a los jurados tomar notas durante el juicio. La justificación era que tomar notas distraía a los jurados, pero ante la complejidad de los juicios contemporáneos, la prohibición obliga a los jurados a recurrir a su memoria, con todos los peligros que ello conlleva. La regla ha ido siendo relajada en las últimas décadas.

[11] Laudan (2006) discute varios ejemplos de las instrucciones que se les da a los jurados en Estados Unidos para ayudarles a entender el estándar de prueba «más allá de toda duda razonable», instrucciones que suelen empeorar su comprensión del estándar.

lideren los procesos de cambio. Una vez se ha creado conciencia sobre el problema, se genera una obligación colectiva de mejorar la situación. Sin embargo, dejar en manos del azar que surjan dichos individuos no parece ser un remedio adecuado para un problema apremiante. Asociado con el argumento anterior, la ausencia de remedios se podría explicar recurriendo al fenómeno psicológico conocido como el «efecto del espectador» (*bystander effect*). En su formulación original, el efecto describe cómo las personas están menos dispuestas a ayudar a alguien que necesita ayuda en una situación de violencia o acoso cuando hay otras personas presentes (Latané y Darley, 1968). Pero el efecto también ocurre cuando un grupo debe completar conjuntamente una tarea; en tales casos, cada individuo tendrá un sentido débil de responsabilidad e intentará hacerle el quite a los problemas que vayan surgiendo. Ante la inacción de los demás, asumirá que el problema es menos importante de lo que parece. Este fenómeno es conocido como la «responsabilidad diluida» (Darley y Latané, 1968). Podemos extender esta forma del efecto del espectador a casos en los que varias personas observan un daño alético que podría ser prevenido o mitigado, pero no intervienen porque asumen que la inacción de los demás refleja la convicción del grupo de que no se requiere ninguna acción (Schwenkenbecher, 2021). Así, paradójicamente un sesgo cognitivo termina impidiendo que alguien actúe para evitar el efecto negativo de los sesgos cognitivos.

Otra posible explicación de la inacción institucional la constituye lo que llamaré la *excusa de la impotencia*. Con respecto a cualquier problema que exija necesariamente una respuesta coordinada de varios agentes, el conocimiento de una solución colectiva clara y definitiva por parte de los agentes es un prerrequisito para la responsabilidad epistémica. Sin embargo, se podría aducir, nadie sabe realmente cómo solucionar el problema de los sesgos cognitivos. Como lo demuestran numerosos estudios, las soluciones que se han planteado para controlar el efecto de los sesgos son pocas y difíciles de implementar (Páez, 2021). Por lo tanto, no hay ninguna obligación de, ni motivación para responder al problema. Es fácil ver que esta respuesta encarna una actitud derrotista frente al asunto. El estudio de los sesgos cognitivos es un área de investigación activa en psicología y no hay por qué pensar que no sea posible identificar normas y prácticas que puedan ser mejoradas. Sin embargo, hay que admitir que históricamente el derecho ha sido indiferente a los descubrimientos de las ciencias cognitivas, como lo han señalado varios psicólogos frustrados ante esta actitud (Saks y Baron, 1980; Monahan y Walker, 1988; Small, 1993). Peor aún, muchos jueces tienen la tendencia a admitir evidencia psicológica de dudosa calidad y existe poco interés en implementar controles de calidad. Un estudio reciente mostró que solo el 40 por ciento de las herramientas de evaluación psicológica utilizadas en las cortes de Estados Unidos durante los años 2016-2018 han sido evaluadas favorablemente por los expertos (Neal *et al.*, 2019).

Finalmente, podemos preguntarnos si en realidad existe una obligación individual o colectiva de tomar acciones para disminuir el riesgo epistémico causado por los sesgos cognitivos. Un defensor del statu quo podría esgrimir lo que llamaré la *excusa de la ignorancia*. Según esta defensa, solo puede haber responsabilidad moral

por un acto u omisión cuando se conocen sus efectos o consecuencias. Ahora bien, los daños aléticos que se derivan de los sesgos cognitivos no son ampliamente conocidos por parte de los agentes judiciales. Por lo tanto, la condición epistémica para la responsabilidad moral no se puede cumplir y no existe responsabilidad alguna por no emprender acciones para prevenir el daño alético. Para analizar esta respuesta, podemos tomar dos perspectivas distintas. Por una parte, desde una perspectiva empírica, la pregunta relevante es si es verdad que no son ampliamente conocidos. Aunque hace veinte años era plausible argumentar que la mayoría de los agentes judiciales no estaban enterados del problema, cada día esa excusa es menos plausible dados los esfuerzos de diversos académicos para crear conciencia sobre el asunto y la popularización del tema a través de libros divulgativos. Por otra parte, desde una perspectiva normativa, podemos preguntarnos si los miembros de la rama judicial *deberían* saber sobre estos daños. Es razonable pensar que existe tanto un deber profesional de educarse a uno mismo como un deber institucional de educar a los jueces y demás agentes de la rama judicial. No hacerlo constituye, desde mi punto de vista, un acto de *negligencia epistémica*, individual y colectiva. En la siguiente sección examinaré cómo se podrían atribuir responsabilidades frente a dicha negligencia.

5. ¿ES POSIBLE ATRIBUIR RESPONSABILIDADES POR LOS DAÑOS GENERADOS POR ESE RIESGO?

Asumiendo que existe algún grado de negligencia epistémica ante el riesgo de los sesgos cognitivos, la pregunta final que quiero plantear es cómo justificar la atribución de responsabilidad a los diferentes agentes que participan en los procesos judiciales[12]. Para tal fin, voy a utilizar el análisis que hacen Fleisher y Šešelja (2020) de las obligaciones epistémicas colectivas. Los autores comienzan planteando la siguiente pregunta: «Si una comunidad científica dada se enfrenta a un daño epistémico que solo puede ser prevenido a través de una acción colectiva, ¿qué clase de deberes epistémicos recaen en cada uno de los científicos?» (p. 1). La pregunta puede ser fácilmente trasladada al ámbito del derecho, donde nos podemos preguntar acerca de los deberes epistémicos de los agentes judiciales frente a la posibilidad de que ocurran daños aléticos causado por los sesgos cognitivos. La propuesta de Fleisher y Šešelja tiene la ventaja adicional de ser una propuesta «prevencionista» (p. 3), es decir, no se enfoca en atribuir responsabilidades por daños causados sino por no implementar normas que prevengan su ocurrencia:

[12] En este artículo solo me ocupo de la responsabilidad que les cabe a los agentes por los daños epistémicos causados por los sesgos *cognitivos*, pero existe una creciente literatura que se ocupa de la responsabilidad individual y colectiva frente a los sesgos sociales como el racismo y el sexismo, comúnmente llamados «sesgos implícitos» (Arena, 2021; Haslanger, 2015; Holroyd, 2012; Holroyd *et al.*, 2017; Rodríguez, 2024; Toribio, 2022). Los sesgos sociales operan psicológicamente de una manera muy diferente a los sesgos cognitivos, y por ende la responsabilidad que se desprende de ellos y los mecanismos para mitigarlos son muy distintos.

> La razón principal por la que las teorías tradicionales de responsabilidad epistémica colectiva necesitan ser suplementadas con una teoría prevencionista es que prevenir una falla al intentar alcanzar una meta epistémica puede requerir normas diferentes a las que típicamente guían la formación de creencias. En otras palabras, un grupo de agentes que están comprometidos con formar sus creencias (o alcanzar conocimiento) de manera responsable, pueden no actuar en la prevención de daños epistémicos latentes (p. 3).

Traducido al ámbito del derecho, el punto de los autores es que podemos entrenar a los jueces en razonamiento probatorio, enseñarles reglas de racionalidad para la argumentación jurídica (Atienza, 1991) y tomar muchas otras medidas que mejoren su proceso de toma de decisiones sobre los hechos, pero mientras no haya una intervención a nivel de las normas procesales para prevenir el riesgo alético causado por los sesgos cognitivos en ese y en otros momentos procesales, la tarea de mejorar el proceso de búsqueda de la verdad se quedará coja.

Examinemos la propuesta en más detalle. Podemos comenzar recordando lo que llamé la «excusa institucional» en el apartado anterior: dado que el poder de un individuo para disminuir el riesgo alético en el derecho es muy limitado, debemos delegar en las instituciones los tipos de cambios procesales que creemos que pueden remediar la situación. Fleisher y Šešelja comienzan admitiendo que hay daños epistémicos que solo pueden ser prevenidos por el esfuerzo conjunto de varios agentes individuales, y nunca por un solo individuo. A este tipo de daño lo denominan *daño epistémico colectivo*[13]. En el caso del derecho, el daño alético generado por los sesgos cognitivos solo puede ser prevenido por la acción conjunta de la gran mayoría de los miembros de la rama judicial. Es decir, es un daño epistémico colectivo.

Esto no significa, sin embargo, que los individuos puedan delegar su responsabilidad en el colectivo o en la institución. Inspirados en Hindriks (2019), los autores introducen una obligación colectiva que llaman la *obligación epistémica de unir fuerzas*. Esta consiste en la obligación inicial de los integrantes de un grupo de unirse para enfrentar el riesgo epistémico, y de actuar posteriormente para prevenirlo. Así, esta obligación tiene dos sub-deberes:

> *Deber 1*: El deber de los agentes de buscar otros agentes para crear conciencia sobre el daño epistémico, expresar su deseo de prevenirlo e invitar a otros a hacer lo mismo.
> *Deber 2*: Las personas que han cumplido con el deber anterior constituyen un *grupo movilizado*, y este tiene el deber de prevenir el daño epistémico.

Así, según esta propuesta, el colectivo formado por los miembros de la rama judicial tiene la obligación epistémica de unir fuerzas para prevenir el daño epistémico colectivo, y una vez movilizados, de actuar, por ejemplo, elevando peticiones urgentes al legislativo para introducir reformas en los códigos de procedimiento. El Deber 2 solo se activa si en efecto se ha constituido un grupo movilizado. Dicho grupo no requiere una estructura. Se trata simplemente de un colectivo con un propósito común. Ahora bien, si no se genera un grupo movilizado que asuma el deber

[13] Lo colectivo no se refiere a cuántas personas son afectadas por el daño sino al tipo de esfuerzo que se requiere para prevenirlo.

de fomentar medidas preventivas, la rama judicial es responsable del daño por haber incumplido la obligación epistémica colectiva de unir fuerzas. Pero esto no excluye las responsabilidades individuales, puesto que cada individuo habrá violado el Deber 1 de crear conciencia sobre el daño epistémico. Por ende, es responsable de no contribuir a la solución conjunta al daño epistémico colectivo.

Como vimos anteriormente, la propuesta de Miranda Fricker es similar a la de Fleisher y Šešelja en el sentido de que el conocimiento colectivo del daño gracias a la labor de algún agente individual genera obligaciones colectivas. Según Fricker,

> El hecho de que los individuos probablemente no puedan hacer mucho actuando solos hace evidente el hecho de que cumplir con las obligaciones individuales en los contextos institucionales tenderá a requerir de una agitación que inspire una respuesta más colectiva hacia la obligación de introducir mejoras. Tales agitaciones a menudo serán suficientes para generar nuevas obligaciones de parte de los cuerpos colectivos relevantes en la institución, en gran medida porque el crear conciencia de las fallas tiende a generar la obligación de mejorar en la medida que sea posible (2016, p. 49).

En otras palabras, cuando un grupo adquiere conciencia de un daño epistémico, es imposible sustraerse a la responsabilidad colectiva de actuar; en tales casos habrá, dice Fricker, «repercusiones normativas colectivas» (p. 49). Ignorar el conocimiento del daño y no actuar constituirá negligencia epistémica colectiva.

Ninguna de las dos propuestas asume que los agentes tienen una posición específica dentro de la jerarquía institucional. De hecho, la propuesta de Fleisher y Šešelja, en particular, está pensada para grupos no organizados. Sin embargo, si queremos aterrizar la propuesta en el contexto del aparato judicial, es evidente que individuos sin poder alguno en la jerarquía judicial difícilmente podrán movilizar a sus colegas y superiores. Los agentes deben poder cumplir con lo que se conoce como la *condición causal* de la responsabilidad (Hindriks, 2019). Aunque es difícil decir exactamente quién cumple con esta condición y por tanto quién debe cumplir el Deber 1, en mi opinión es un deber que recae, por lo menos, sobre todos los jueces y todos los miembros de órganos superiores en la jerarquía judicial, y quizás en fiscales, procuradores y cargos semejantes. Es más probable, además, que personas mejor ubicadas en la jerarquía hayan tenido la oportunidad de educarse sobre los riesgos aléticos generados por los sesgos. Es decir, son personas que cumplen con la *condición epistémica* para la responsabilidad: «Los agentes tienen una creencia justificada acerca del daño inminente y sobre la probabilidad de tener éxito en su prevención (potencialmente, a través del esfuerzo colectivo)» (Fleisher y Šešelja, 2020, p. 6).

La obligación epistémica de unir fuerzas no es absoluta. La propuesta deja espacio para que los agentes puedan argumentar que existe un conflicto de obligaciones que les impide cumplir con el Deber 1. El agente puede aducir que debe cumplir con obligaciones que considera superiores, como las descritas al comienzo del apartado anterior. El temor, por ejemplo, a que la consecuencia de crear un grupo movilizado sea poner en riesgo los derechos de los imputados debe ser tenido en consideración al evaluar la responsabilidad epistémica de cualquier agente.

Finalmente, Lerner y Tetlock (2003) presentan evidencia experimental de que, bajo ciertas circunstancias, las personas que asumen responsabilidades mejoran sus juicios y decisiones, en el sentido de que estas se verán menos afectadas por sesgos cognitivos. Esto sugiere que la introducción de una cultura de la responsabilidad manifiesta (*accountability*) en la rama judicial puede contribuir a atenuar el riesgo alético causado por los sesgos cognitivos. Los efectos positivos de la responsabilidad individual se pueden ver reforzados si se genera un grupo movilizado en donde todos los miembros del grupo asumen la responsabilidad colectiva de prevenir los daños epistémicos.

6. CONCLUSIONES

En este artículo he presentado una mirada amplia al problema del riesgo epistémico, más allá de los simples errores tipo I y II que han dominado las discusiones sobre el error en el derecho. Adoptar esta visión ampliada del riesgo epistémico, que examina el proceso completo de búsqueda de la verdad y no solo el paso final en el que se acepta la conclusión, tiene varias ventajas. Por una parte, nos permite ver que el daño epistémico causado por los sesgos cognitivos puede ocurrir en muchos momentos diferentes de un proceso judicial, no solo en el momento de la decisión final, y que estos pueden afectar a muchos agentes diferentes. Tanto en la generación de pruebas en la etapa de investigación, como en la presentación de estas en el juicio, existen elementos no epistémicos que deben ser sumados al riesgo de error. Además, extiende la responsabilidad por el daño a los diferentes agentes judiciales, no solo al decisor de los hechos, y plantea la pregunta de si estos puedan ser responsabilizados individualmente por no tomar acciones para prevenir esos daños. Como vimos en la sección final, hay formas de atribuir responsabilidades tanto individuales como colectivas cuando se trata de daños epistémicos. Finalmente, permite ver que puede haber varios valores no epistémicos que pueden intervenir en la forma en que se enfrenta el riesgo epistémico de los sesgos cognitivos. La discusión de cualquier esfuerzo para modificar las normas procesales para disminuir el riesgo epistémico tendrá que comenzar necesariamente por discutir esos valores no epistémicos, como una preferencia marcada por preservar el *stare decisis* y el principio de inmediación. La mala noticia es que, si la introducción de mejores procesos de decisión pasa por la discusión de esos principios, sería ingenuo pensar que las reformas van a ser introducidas fácilmente en el futuro cercano. La única manera de introducir un sentido de urgencia es introducir evaluaciones de la calidad de las decisiones pasadas para revelar sus falencias, y continuar con la educación judicial sobre los sesgos cognitivos.

REFERENCIAS

Arcila-Valenzuela, M., y Páez, A. (2024). Testimonial injustice: The facts of the matter. *Review of Philosophy and Psychology*, *15*(2), 585-602.

Arena, F. (2021), Responsabilidad por sesgos implícitos y decisión judicial. En F. Carbonell & J. Valenzuela (Eds.), *Fundamentos filosóficos del derecho procesal* (pp. 47-80). Tirant Lo Blanch.

Atienza, M. (1991). *Las razones del derecho. Teorías de la argumentación jurídica*. Centro de Estudios Constitucionales.

Austin, W., y Williams III, T. A. (1977). A survey of judges' responses to simulated legal cases: Research note on sentencing disparity. *Journal of Criminal Law & Criminology*, *68*(2), 306-310.

Biddle, J. B. (2007). Lessons from the Vioxx debacle: what the privatization of science can teach us about social epistemology. *Social Epistemology, 21*(1), 21-39.

Biddle, J. B., y Kukla, R. (2017). The geography of epistemic risk. En K. C. Elliott & T. Richards (Eds.), *Exploring inductive risk: Case studies of values in science* (pp. 215-237). Oxford University Press.

Bordalo, P., Coffman, K., Gennaioli, N., y Shleifer, A. (2016). Stereotypes. *The Quarterly Journal of Economics*, *131*(4), 1753-1794.

Brownstein, M. (2018). *The implicit mind: Cognitive architecture, the self, and ethics*. Oxford University Press.

Brownstein, M., y Saul, J. (Eds.). (2016a). *Implicit bias and philosophy, volume 1: Metaphysics and epistemology*. Oxford University Press.

Brownstein, M., y Saul, J. (Eds.). (2016b). *Implicit bias and philosophy, volume 2: Moral responsibility, structural injustice, and ethics*. Oxford University Press.

Buller, D. B., y Burgoon, J. K. (1996). Interpersonal deception theory. *Communication Theory*, *6*(3), 203-242.

Clancy, K., Bartolomeo, J., Richardson, D., y Wellford, C. (1981). Sentence decisionmaking: The logic of sentence decisions and the extent and sources of sentence disparity. *Journal of Criminal Law & Criminology*, *72*(2), 524-554.

Collins, A. W. (1996). Moore's Paradox and epistemic risk. *Philosophical Quarterly*, *46*(148), 308-319.

Coloma, R., Larroucau, J., y Páez, A. (2024). Sobre el impacto judicial de la concepción racionalista de la prueba. *Revus. Journal for Constitutional Theory and Philosophy of Law*, *53*. DOI: 10.4000/123oc.

Darley, J. M., y Latané, B. (1968). Bystander intervention in emergencies: diffusion of responsibility. *Journal of Personality and Social Psychology*, *8*(4), 377-383.

Douglas, H. (2000). Inductive risk and values in science. *Philosophy of Science*, *67*(4), 559-579.

Douglas, H. (2017). Why inductive risk requires values in science. En K. C. Elliott & D. Steel (Eds.), *Current controversies in values and science* (pp. 81-93). Routledge.

Fischhoff, B., Slovic, P., y Lichtenstein, S. (1978). Fault trees: Sensitivity of estimated failure probabilities to problem representation. *Journal of Experimental Psychology: Human Perception and Performance*, *4*, 330-344.

Fleisher, W., y Šešelja, D. (2020). Collective epistemic responsibility: a preventionist account. https://philsci-archive.pitt.edu/17003/.

Frankel, M. E. (1973). *Criminal sentences: Law without order*. Hill and Wang.

Fricker, M. (2016). Fault and no-fault responsibility for implicit prejudice: A space for epistemic 'agent-regret'. En M. S. Brady & M. Fricker (Eds.), *The epistemic life of groups* (pp. 33-50). Oxford University Press.

Gigerenzer, G. (2006). Heuristics. En G. Gigerenzer & C. Engel (Eds.), *Heuristics and the law* (pp. 17-44). MIT Press.

Gillis, N. B. (2021). Sexism in the judiciary: The importance of bias definition in NLP and in our courts. En *Proceedings of the 3rd Workshop on Gender Bias in Natural Language Processing* (pp. 45-54). ACM.

Haslanger, S. (2015), Distinguished lecture: Social structure, narrative, and explanation. *Canadian Journal of Philosophy*, *45*(1), 1-15.

Hempel, C. G. (1965). Science and human values. En *Aspects of scientific explanation* (pp. 81-96). The Free Press.

Hindriks, F. (2019). The duty to join forces: When individuals lack control. *The Monist*, *102*(2), 204-220.

Hirsch, A. J. (2004). Evolutionary theories of common law efficiency: Reasons for (cognitive) skepticism. *Florida. State University Law Review*, *32*(2), 425-441.

Holroyd, J. (2012). Responsibility for implicit bias. *Journal of Social Philosophy*, *43*(3), 274-306.

Holroyd, J., Scaife, R., y Stafford, T. (2017). Responsibility for implicit bias. *Philosophy Compass*, *12*(3), e12410.

Jois, G. U. (2009). Stare decisis is cognitive error. *Brooklyn Law Review*, *75*(1), 63-141.

Kahneman, D., Slovic, P., y Tversky, A (Eds.) (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge University Press.

Kahneman, D., y Tversky, A. (1973). On the psychology of prediction. *Psychological Review*, 80, 237-251.

Kuran, T., y Sunstein, C. R. (1998). Availability cascades and risk regulation. *Stanford Law Review*, *51*, 683-768.

Latané, B., y Darley, J. M. (1968). Group inhibition of bystander intervention in emergencies. *Journal of Personality and Social Psychology*, *10(3)*, 215-221.

Laudan, L. (2006). *Truth, error, and criminal law: an essay in legal epistemology*. Cambridge University Press.

Lerner, J. S., y Tetlock, P. E. (2003). Bridging individual, interpersonal, and institutional approaches to judgment and decision making: The impact of accountability on cognitive bias. En S. L. Schneider & J. Shanteau (Eds.), *Emerging perspectives on judgment and decision research* (pp. 431-457). Cambridge University Press.

Lidén, M. (2023). *Confirmation bias in criminal cases*. Oxford University Press.

Marshall, B. C., y Alison, L. J. (2007). Stereotyping, congruence and presentation order: Interpretative biases in utilizing offender profiles. *Psychology, Crime & Law*, *13*(3), 285-303.

Miller, N., y Campbell, D. T. (1959). Recency and primacy in persuasion as a function of the timing of speeches and measurements. *The Journal of Abnormal and Social Psychology*, *59*(1), 1-9.

Monahan, J., y Walker, L. (1988). Social science research in law: A new paradigm. *American Psychologist*, *43*, 465-472.

Myers, D. G., y Lamm, H. (1976). The group polarization phenomenon. *Psychological Bulletin*, *83*, 602-627.

Neal, T. M., Slobogin, C., Saks, M. J., Faigman, D. L., y Geisinger, K. F. (2019). Psychological assessments in legal contexts: Are courts keeping «junk science» out of the courtroom?. *Psychological Science in the Public Interest*, *20*(3), 135-164.

Nisbett, R. E., y Wilson, T. D. (1977). The halo effect: Evidence for unconscious alteration of judgments. *Journal of Personality and Social Psychology*, *35*(4), 250-256.

O'Brien, B., y Grosso, C. M. (2020). Criminal trials and reforms intended to reduce the impact of race: A review. *Annual Review of Law and Social Science*, *16*(1), 117-130.

Páez, A. (2021). Los sesgos cognitivos y la legitimidad racional de las decisiones judiciales. En F. J. Arena, P. Luque, y D. Moreno Cruz (Eds.), *Razonamiento jurídico y ciencias cognitivas* (pp. 187-222). Universidad Externado de Colombia.

Páez, A., y Matida, J. (2023). La injusticia epistémica en el proceso penal. *Milan Law Review*, *4*(2), 114-136.

Partridge, A., y Eldridge, W. B. (1974). *The Second Circuit sentencing study: A report to the judges of the Second Circuit*. Federal Judicial Center.

Posner, R. A. (1999). An economic approach to the law of evidence. *Stanford Law Review*, *51*, 1477-1546.

Rachlinski, J. J. (1998). A positive psychological theory of judging in hindsight. *The University of Chicago Law Review*, *65*, 571-625.

Rasmussen, N. (1975). Reactor safety study. *WASH-1400*. US NRC.

Rodríguez, H. A. (2024). Sesgos implícitos, injusticia explícita: Efectos epistémicos de los sesgos inconscientes en el razonamiento probatorio en México. *Quaestio Facti*, 7, 103–135.

Ross, L. (1977). The intuitive psychologist and his shortcomings: Distortions in the attribution process. En L. Berkowitz (Ed.), *Advances in Experimental Social Psychology. Vol. 10* (pp. 173–220). Academic Press.

Saks, M. J., & Kidd, R. F. (1980). Human information processing and adjudication: Trial by heuristics. *Law and Society Review*, *15*, 123-160.

Saks, M. J., & Baron, C. H. (Eds.) (1980). *The use/nonuse/misuse of applied social research in the courts*. Abt Books.

Schauer, F. (2006). On the supposed jury-dependence of evidence law. *University of Pennsylvania Law Review*, *155*, 165-202.

Schwenkenbecher, A. (2021). Structural injustice and massively shared obligations. *Journal of Applied Philosophy*, *38*(1), 23-39.

Small, M. A. (1993). Legal psychology and therapeutic jurisprudence. *St. Louis University Law Review*, *37*, 675-713.

Toribio, J. (2022). Responsibility for implicitly biased behavior: A habit-based approach. *Journal of Social Philosophy*, *53*(2), 239-254.

Tuerkheimer, D. (2017). Incredible women: Sexual violence and the credibility discount. *University of Pennsylvania Law Review*, *166*(1), 1-58.

Weinstein, I. (2003). Don't believe everything you think: Cognitive bias in legal decision making. *Clinical Law Review, 9*, 783-834.

Wellborn III, O. G. (1991). Demeanor. *Cornell Law Review*, *76*(5), 1075-1105.

Wilholt, T. (2009). Bias and values in scientific research. *Studies in History and Philosophy of Science, Part A*, *40*(1), 92– 101.

Quaestio facti. Revista Internacional sobre Razonamiento Probatorio
Quaestio facti. International Journal on Evidential Legal Reasoning
Sección: Homenaje a Larry Laudan
2026 I 10 pp. 67-94
Madrid, 2025
DOI: 10.33115/udg_bib/qf.i10.23126
Marcial Pons Ediciones Jurídicas y Sociales
© Rodrigo Coloma Correa
© Renato Lira Rodríguez
© Sebastián Sandoval Ayala
ISSN: 2604-6202
Recibido: 09/03/2025 | Aceptado: 19/11/2025 | Publicado *online*: 10/12/2025
Editado bajo licencia Reconocimiento 4.0 Internacional de Creative Commons

DUDAS RAZONABLES. SOBRE LA CONCEPCIÓN DE ESTÁNDAR DE PRUEBA DE LARRY LAUDAN*

Rodrigo Coloma Correa**
Renato Lira Rodríguez***
Sebastián Sandoval Ayala****

RESUMEN: En este trabajo discutimos dos tesis y un presupuesto de la concepción de Larry Laudan sobre los estándares de prueba. Se trata de que: (1) los estándares subjetivos y/o ambiguos no son estándares en el sentido estricto de la expresión; (2) los estándares de prueba son mecanismos diseñados para distribuir falsos positivos y falsos negativos; y (3) los estándares de prueba se reconstruyen adecuadamente como umbrales. Concluimos que la primera afirmación no se encuentra justificada (los estándares de prueba pueden ser subjetivos y ambiguos), la segunda es de escasa utilidad (lo indeterminable no puede ser distribuido), y la tercera resulta reduccionista (la metáfora del umbral deja de lado una forma importante en que operan los estándares de prueba).

* La contribución de cada uno de los autores ha sido equivalente. Agradecemos las observaciones de quienes evaluaron anónimamente esta investigación. El trabajo se ha elaborado en el marco de Imputatio: *Centro de Análisis sobre la atribución de intenciones y la imputación de responsabilidades* (www.imputatio.cl). El segundo autor agradece el patrocinio de ANID-Subdirección de Capital Humano/Doctorado Nacional/2023-21230033 y del Programa de Doctorado en Derecho, mención constitucionalismo y derecho, de la Universidad Austral de Chile.

** Profesor titular de la Universidad Alberto Hurtado de Chile. https://orcid.org/0000-0003-3347-7625.

*** Doctorando en Derecho, mención Constitucionalismo y Derecho, Universidad Austral de Chile. Máster en Estado de Derecho y Democracia Constitucional por la Universidad de Génova. https://orcid.org/0000-0003-1119-0834.

**** Doctorando en Derecho, Universidad Alberto Hurtado. Magíster en Derecho, mención en Derecho Público, de la Universidad de Chile. Profesor del Departamento de Teoría y Filosofía del Derecho de la Universidad Diego Portales. https://orcid.org/0000-0002-9355-5907.

Palabras clave: estándar de prueba; Larry Laudan; error; prototipo; Prueba.

REASONABLE DOUBTS. ON LARRY LAUDAN'S CONCEPTION OF THE STANDARD OF PROOF

ABSTRACT: In this paper, we discuss two theses and a presupposition of Larry Laudan's conception of standards of proof. They are that: (1) subjective and/or ambiguous standards are not standards in the strict sense of the term; (2) standards of proof are mechanisms designed to distribute false positives and false negatives; and (3) standards of proof are adequately reconstructed as thresholds. We conclude that the first claim is not justified (standards of proof can be subjective and ambiguous), the second is of limited utility (the indeterminable cannot be distributed), and the third is reductionist (the threshold metaphor overlooks an important way in which standards of proof operate).

Keywords: standard of proof; Larry Laudan; Error; prototype; evidence.

SUMARIO: 1. INTRODUCCIÓN.— 2. LOS ESTÁNDARES DE PRUEBA SUBJETIVOS Y/O AMBIGUOS SÍ PUEDEN SER ESTÁNDARES.— 3. LOS EDP NO SIRVEN PARA DISTRIBUIR FALSOS POSITIVOS Y FALSOS NEGATIVOS.— 4. LOS EDP NO (SOLO) SON RECONSTRUIBLES COMO UMBRALES.— 5. CONCLUSIÓN.— BIBLIOGRAFÍA.

«¿Quién dijo que todo está perdido?»
Fito Páez, *Yo vengo a ofrecer mi corazón*

1. INTRODUCCIÓN

En el campo de la teoría de la prueba se discute intensamente sobre los estándares de prueba (en adelante, EdP). Las controversias versan sobre numerosos aspectos, como la posibilidad de fijarlos, la conveniencia de hacerlo, el grado de exigencia probatoria adecuado para tomar decisiones, o la forma en que se expresa la suficiencia probatoria en los textos normativos. Al respecto, la influencia de Larry Laudan es indiscutible. Sus investigaciones sobre el EdP focalizadas en el «más allá de toda duda razonable» (MADR) —extensibles a otros estándares probatorios y difundidas, principalmente, en «Por qué un estándar de prueba subjetivo y ambiguo no es un estándar» y en «Verdad, Error y Proceso Penal»— instalaron entre nosotros un conjunto de ideas que, hasta hoy, marcan la manera en que entendemos los EdP.

En este trabajo pondremos en duda dos tesis de la concepción de Laudan y uno de sus presupuestos, a saber: (1) que los estándares subjetivos y/o ambiguos no son estándares en el sentido estricto de la expresión (tesis); (2) que su función principal es la distribución de falsos positivos y falsos negativos (tesis); y (3) que es adecuado reconstruirlos como umbrales (presupuesto). A nuestro juicio, la primera afirmación no está justificada, la segunda es de escasa utilidad, y la tercera es reduccionista si se tiene en cuenta cómo se desarrolla la práctica judicial.

2. LOS ESTÁNDARES DE PRUEBA SUBJETIVOS Y/O AMBIGUOS SÍ PUEDEN SER ESTÁNDARES

Las preocupaciones de Laudan al estudiar el funcionamiento de los sistemas judiciales y, más específicamente, la manera en que el procedimiento penal define la culpabilidad de los acusados apunta a: i) producir conocimientos acerca de cuán bien o mal determinan la verdad en los asuntos que se investigan; y ii) proponer mejoras de las prácticas investigativas con miras a reducir los errores que usualmente arrojan. En lo que a los EdP concierne, Laudan (2005, p. 96) pretende aclarar ¿qué son?, ¿cómo funcionan? y ¿a qué se parecen? Para llevar a cabo dicha tarea, recurre a saberes asociados a su destacadísima trayectoria científica, esta es, la de un físico teórico ocupado de cuestiones de epistemología. No es de extrañar, entonces, encontrar en sus textos afirmaciones tales como:

> No me andaré con rodeos, esto [el proceso penal] es una parodia de sistema de prueba. Un EdP —en cualquier ámbito fuera del derecho en que se reclamen las pruebas (incluyendo las ciencias naturales, los ensayos clínicos en la medicina, las matemáticas, los estudios epidemiológicos, etcétera)— tienen la intención de indicar al investigador o aquel que se está cuestionando cuándo está autorizado a considerar como probado, esto es, cuándo la relación entre la prueba o las premisas justifica la aceptación de la conclusión como probada para los propósitos pretendidos (2005, p. 104).

Laudan sostiene que un EdP cuya aplicación en un proceso judicial dependa de los juicios subjetivos[1] del decisor no resulta apto para establecer niveles adecuados de exigencia probatoria[2]. De hecho, indica que ese defecto incide en que el sistema procesal carezca de fiabilidad (entendida como uniformidad y predictibilidad) y sea «incoherentemente injusto» (2013, pp. 62-63). A su entender, si los estados mentales

[1] Como bien observa González Lagier (2020, p. 91), la subjetividad que preocupa a Laudan puede ser leída de dos formas distintas: como una apelación directa a los estados mentales del juzgador (así ocurre, por ejemplo, con fórmulas como íntima convicción) y como sinónimo de vago o impreciso «porque, al ser impreciso (un EdP), acaba decidiendo el juez discrecionalmente, de acuerdo con su criterio subjetivo». La tesis de Laudan asume que podemos eliminar (o al menos reducir significativamente) la vaguedad que afecta a los EdP y, de esta forma, eliminar (o al menos reducir significativamente) la discrecionalidad.

[2] Para Laudan (2005, pp. 100-106), depurar los EdP por la vía de recurrir al cálculo bayesiano, conduce a un subjetivismo no muy diferente al de aquellos EdP que se basan en grados de confianza en los decisores.

de los decisores no están guiados por criterios que fomenten una conexión inferencial entre las pruebas y la decisión, «son propensos a ser infundados, prejuiciosos e irracionales» (2013, p. 90). Por esto, sostiene que, mientras las definiciones de la duda razonable permanezcan obsesionadas con la intensidad de las creencias, estaremos expuestos a la «devastadora crítica de que confunde la fuerza o intensidad de una creencia (que puede ser totalmente irracional) con que la creencia en cuestión esté *justificada*» (2013, pp. 101-102).

Para Laudan, la intensidad de la convicción no es un criterio relevante para determinar una conexión fundada con las pruebas. La dependencia de un EdP en el nivel de confianza del decisor «no aporta nada al proyecto de promocionar la emisión de veredictos válidos» (2013, p. 102)[3]. Así, sostiene que formulaciones como la «convicción estable en la culpabilidad del acusado» o alcanzar un «estado subjetivo cercano a la certidumbre» no son auténticos EdP. Para ilustrar la imposibilidad, piénsese en un caso cualquiera donde A considera que ha alcanzado una convicción estable de la culpabilidad de X, mientras que B entiende que no ha alcanzado tal estado. Acorde a ello, A condenará y B absolverá. ¿Podría decirse que ambos sujetos han aplicado el EdP exigible? Laudan lo pone en duda.

Según Laudan, entonces, el EdP debe ser objetivo, es decir, estar formulado en «términos del vínculo inferencial que debe existir entre las pruebas disponibles y la hipótesis como una proposición probada» (2013, p. 126). Solo de ese modo es posible definir las propiedades o criterios epistémicos que debe cumplir la evidencia (como la refutación de hipótesis alternativas o la confirmación de ciertos datos). Así, su idea «es que las enunciaciones del estándar dejen de ser enunciaciones de principio, y que pasen a tener una relación directa con la prueba de los hechos en el caso particular» (Taruffo, 2005, p. 120)[4].

El problema de la subjetividad, según Laudan (2013, pp. 66-83), se ve favorecido por la ambigüedad de los EdP[5], como resulta palmario en el caso del MADR[6]. En

---

[3] Hay una cierta distancia entre el concepto de validez manejado por Laudan (2013, pp. 33-37) y la mayoría de los juristas. La propuesta de Laudan no parece dar mayor espacio a la idea de que las normas procedimentales tienen un carácter de normas constitutivas.

[4] En la misma línea, Ferrer (2025, p. 178 y ss.) indica que si los EdP no satisfacen las expectativas epistémicas que debe cumplir la prueba, resultará imposible cumplir la obligación de justificar la superación del umbral de suficiencia probatoria. Solo en el caso que aquello sea logrado, será posible el control de la corrección del razonamiento probatorio mediante los recursos, como también la regulación de la manera en que los jueces y tribunales deben ejercer su competencia para decidir sobre los hechos del caso.

[5] Ferrer (2007, p. 152) también destaca la importancia de definir con claridad los EDP: «sin ellos no puede pretenderse una valoración racional de la prueba ni un control de la valoración realizada». Ferrer (2025) sigue a Schauer (2004, pp. 195, 209-221; 2013, p. 201) para sostener que los tribunales deben enfocarse «en la presencia o ausencia de ciertos hechos y permitiéndoles «relajarse» con respecto a otros». Ello impide que «decisores sabios y sensibles tomen las mejores decisiones, a fin de impedir que decisores incompetentes o simplemente injustos tomen decisiones incorrectas».

[6] En el mismo sentido, Accatino (2011, pp. 496-498).

sus palabras: «la idea de que el significado de una duda razonable debe ser claro para todos, por el hecho de que la gente comprende cada uno de los términos que conforman la frase, es hilarante» (2013, p. 84). A lo más, sirve como una mera exhortación a no tomar el encargo decisorio a la ligera, pero no otorga «pistas mínimas acerca de cómo deberían lucir unas pruebas contundentes de culpabilidad» (2013, p. 92). Por esto, Ferrer (2025, p. 189 nota 40) también indica que «de nada sirve adherir el adjetivo «razonable» a la duda porque seguiría siendo condición necesaria del cumplimiento del estándar que el juzgador tenga el estado mental de duda (…)». Es así como Laudan (2005, p. 111) afirma que los genuinos EdP «no dejan mucho margen de ambigüedad respecto a si han sido satisfechos» (2005, p. 111), de modo que los decisores no debieran tener desacuerdos sobre sus obligaciones derivadas del EdP.

Veamos, entonces, cuáles serían para Laudan (2005, p. 107) algunos ejemplares de genuinos EdP. De ellos se espera que la judicatura solo verifique la presencia de criterios epistémicos que debe cumplir la evidencia para sostener que un hecho está probado[7]:

> Si es creíble la prueba acusatoria o un testimonio que resultaría difícil de explicar si el acusado fuera inocente y no es creíble la prueba exculpatoria o un testimonio que sería muy difícil de explicar si el acusado fuese culpable, entonces condénelo. De otro modo, absuélvalo; o bien
>
> Si la historia de la acusación acerca del delito es plausible y usted no puede imaginar una historia plausible que muestre al acusado como inocente, entonces condénelo. De otro modo, absuélvalo.

Hasta aquí, podría decirse que las preocupaciones de Laudan son —¡qué duda cabe!— del todo atendibles: *un EdP que produzca decisiones que no generen discrepancias en una comunidad racional (en cuanto a que se han establecido conexiones sólidas entre la prueba rendida y los hechos probados) es preferible a uno que no vaya en dicha dirección*. Pero ello no implica que la objetividad deba considerarse como una propiedad definitoria de los EdP. Las objeciones hacia ello se producen en dos niveles: uno conceptual y otro epistémico.

La objeción conceptual argumenta que es insostenible plantear a la objetividad como una propiedad definitoria de los EdP. Para justificar nuestro punto de vista, nos vemos forzados a hacer algunas precisiones sobre lo que se quiere decir con la partícula *estándar* y, por extensión, a lo que se dice con la locución *estándar de prueba*. Si bien nuestras formas de hablar y el análisis conceptual no parecen estar dentro de las preocupaciones de Laudan, la forma de construir su argumentación y de titular sus textos hacen ineludible llevar a cabo un modesto ejercicio de depuración.

El Diccionario de la Lengua Española define la palabra *estándar* como: «Que sirve como tipo, modelo, norma, patrón o referencia». A su vez, al tratar la palabra *estan-*

---

[7] Para su formulación tiene en consideración las propuestas de Ronald Allen. Más detalles en nota 9 de Laudan, (2005, p. 107). En forma similar, Ferrer (2025, pp. 57-58) sostiene que, para evitar la contaminación de la decisión probatoria con valores privados (incompatible con el Estado de derecho y el derecho a la prueba) es esencial establecer EdP que garanticen la no arbitrariedad enfocados en la capacidad justificativa y propiedades epistémicas concretas.

darizar, se quiere decir: «tipificar (ajustar a un tipo o norma)». Por cierto, al aludir a este diccionario, no pretendemos que el significado de la noción «estándar» dependa completamente del lenguaje ordinario, ni menos de lo que dice la Real Academia Española. Solo se trata de que, al no haber un desarrollo generalizado *sobre lo que significa esta partícula*, aludir a este recurso facilita la discusión[8]. Más aún, creemos que Laudan no tendría ninguna objeción con el uso de estas definiciones, pues sus esfuerzos al tratar los EdP apuntan precisamente a reducir la dispersión en las decisiones asegurando un cierto tipo de relación[9]. Así las cosas, si un EdP subjetivo y/o ambiguo provocase un mayor alineamiento en las decisiones probatorias, sería un candidato a EdP.

Teniendo en cuenta lo recién señalado, prestemos atención a lo que nos dice Laudan acerca de lo que ocurre si se aplica una formulación del MADR, en clave subjetiva y ambigua. En una nota de Verdad, Error y Proceso Penal (2013, nota 5, p. 112) sostiene:

> Lo que Kalven y Zeisel encontraron, no fue que dos de cada 10 absoluciones correspondían a personas que fueran culpables, sino que dos de cada diez absoluciones, en la opinión de los jueces, ameritaban una condena justificada; o sea que esas dos absoluciones tenían todo para satisfacer el EdP en materia penal. Dado que dicho estándar es bastante severo, tenemos buenas razones para esperar que las absoluciones falsas, del modo en que venimos empleando la expresión —es decir, la absolución de acusados genuinamente culpables— ocurran con mucho mayor frecuencia de lo que la cifra 20 por 100 podría implicar.

Desde el punto de vista del uso y significado de las palabras, lo recién observado no pone en tela de juicio que los EdP subjetivos y ambiguos sean ejemplares de la clase genérica de los estándares. Y esto es así porque hay un 80 por 100 de absoluciones en que se producen coincidencias entre jurados y jueces. El porcentaje no es bajo considerando que jurados y jueces pertenecen a comunidades disciplinares diferentes. Así, a nivel conceptual, podríamos decir que el EdP *ha estandarizado* las decisiones. Aunque no ha estandarizado de la manera en que ético-políticamente nos gustaría, ello no justifica privar a formulaciones calificadas como subjetivas y ambiguas del estatus de EdP.

A la luz de lo señalado, alguien podría decir: *De acuerdo, un EdP subjetivo y ambiguo estandariza, pero no como cabe esperar de un EdP*. En otras palabras, si los jueces condenan solo si tienen una «convicción estable en la culpabilidad del acusado», la estandarización podría lograrse sin necesidad de que dicha convicción tenga que ver con un vínculo inferencial entre las pruebas rendidas y los hechos que se dan por probados. Así, por ejemplo, un juez racista podría formar una convicción estable de culpabilidad cuando el acusado pertenezca a cierta etnia, y aquello, por cierto,

---

[8] Sobre el punto: Gama, (2016, pp. 61 y ss.) y Coloma, (2016a; 2016b, pp. 108 y ss). En un sentido opuesto, Aguilera, (2023).

[9] La dispersión se asocia a lo que más adelante categorizamos como «ruido» para hablar sobre la variabilidad de las decisiones humanas en asuntos donde se espera mayor coincidencia (Kahneman, Sibony y Sunstein, 2021, p. 42).

no se entendería como una forma de conectar las pruebas rendidas con los hechos probados. Lo mismo podría ocurrir con un conjunto de jueces indolentes que simplemente hagan depender su decisión del azar. En casos así, por cierto, no se estaría aplicando un EdP sino otra forma de tomar decisiones que aquí no interesa.

La idea de que un EdP formulado en términos subjetivos aún sirva para estandarizar se refuerza gracias a la circunstancia de que el EdP funciona de manera coordinada con otras reglas. Este sería el caso de la obligación de motivar las decisiones judiciales, como también, la de llevar a cabo un ejercicio de deliberación si se forma parte de un tribunal colegiado[10]. Así las cosas, el tribunal deberá superar el filtro de la deliberación y/o de la motivación que lo llevarán a explicitar razones que tienen que ver con la conexión entre las pruebas rendidas y los hechos probados. Si así no ocurre, el sistema prevé que las decisiones sean modificadas y/o que el juez pueda ser sancionado. Estas son las mismas garantías que operarían para controlar a tribunales que, debiendo aplicar un EdP objetivo (suponiendo que es posible/conveniente formularlos), no lo hagan. De esta manera, en la medida que los sistemas procesales consideren algún nivel de control, también un EdP subjetivo y con ciertos niveles de ambigüedad sería conceptualmente un EdP.

Vamos ahora a la objeción epistémica sobre los EdP subjetivos y/o ambiguos. Lo que hemos dicho hasta el momento no se hace cargo de la crítica de que los EdP subjetivos y/o ambiguos podrían no ser los más útiles para que nuestros procedimientos alcancen un nivel adecuado de corrección en la distribución de cargas y/o beneficios. Laudan tiene un punto indudable a su favor y que, probablemente, ha sido un importante factor del éxito de su teoría: *las formulaciones de los EdP no pueden desentenderse de que, para la decisión de dar por probados los hechos en un proceso judicial, es clave la calidad (robustez) de la conexión que se construye entre ellos y la prueba rendida*. En lo que sigue, no negamos que se trata de una pretensión relevante y que debe estar presente en cualquier intento de depuración de los EdP. El problema es que la propuesta de Laudan pareciera desconocer que, en un proceso judicial, la toma de decisiones probatorias está vinculada a la toma de decisiones interpretativas y valorativas. Como diagnóstico, obsérvese que los candidatos propuestos por Laudan como EdP genuinos utilizan palabras tales como: «creíble», «difícil de explicar», «plausible» o «usted no puede imaginar». Ninguna de ellas parece erradicar la subjetividad del decisor[11].

Revisemos esta idea más en detalle. Laudan advierte una conexión fuerte entre la equivocidad que afecta a las disposiciones que expresan los EdP y la relevancia que

[10] En descargo de Laudan podría señalarse que él se enfoca primordialmente en sistemas anglosajones en que opera un jurado que no está sujeto a la obligación de motivar los hechos probados, aunque sí de deliberar. Sin embargo, el propio Laudan (2005, p. 98) se preocupa de hacer algunos guiños al sistema continental.

[11] La crítica que formulamos a continuación no es del todo novedosa. Los candidatos propuestos por Laudan para asumir el papel de un EdP objetivo no erradican la vaguedad que afecta a las disposiciones normativas ni tampoco se elimina el recurso a la subjetividad al momento de tomar una decisión.

cobra el juzgador al momento de tomar una decisión. La intuición parece indicar que, al disminuir la equivocidad que afecta a los textos normativos, también se disminuye el espacio para la subjetividad. De ahí que sus propuestas busquen robustecer las disposiciones normativas enunciando operaciones que, *prima facie*, serían fácilmente identificables y replicables. Al respecto, nuestra crítica sostiene (1) que los términos a los que Laudan se ve obligado a recurrir para ensayar formulaciones de EdP genuinos («creíble», «difícil de explicar», «plausible» o «usted no puede imaginar») no cumplen con el propósito de eliminar la subjetividad (y equivocidad) de los EdP; y (2) que tales vínculos con lo que es propio del juzgador parece ser un *componente necesario* para adoptar decisiones probatorias.

Vamos con el primer punto. A nuestro entender, la formulación de un EdP siguiendo las recomendaciones de Laudan no altera de manera significativa el hecho de que la suficiencia probatoria siga dependiendo, en última instancia, de la estimación interpretativa y valorativa que realiza el tribunal. Esto ocurre porque las propuestas normativas de EdP recurren a criterios que, en palabras de Dei Vecchi (2020a, p. 41) «son tan graduales como la justificación epistémica misma a la que conforman». Ello provoca que sean los mismos parámetros epistémicos («credibilidad», «plausibilidad», entre otros) los que imponen la necesidad de efectuar juicios de valor y alejarse de su mera constatación [12]. Por ejemplo, como indica Calderón (2023, p. 159) «sostener que una hipótesis resulta plausible es similar a decir que una exhibición resulta obscena: obedece a criterios valorativos que se predican respecto del acto, que dependen de otras variables además de las condiciones o características del acto en sí mismo». Lo interesante, entonces, es que dicha gradualidad no puede entenderse como un defecto lingüístico pronto a ser corregido, sino como una pieza constitutiva del razonamiento probatorio, que nos habilita a realizar juicios contextuales en términos de «más o menos» [13].

Conforme a lo anterior, la toma de decisiones probatorias requiere que el tribunal no solo determine el significado de criterios epistémicos equívocos, sino también que establezca si la evidencia cumple con expectativas de suficiencia. En este sentido, cualquier propuesta sobre EdP debe reconocer que las razones epistémicas están vinculadas al establecimiento discrecional de la suficiencia probatoria. De ahí que la aplicación de un EdP implique elaborar juicios contextuales (asumir que la suficiencia

[12] Esta evaluación es posible precisamente *porque* los términos son vagos tanto en un sentido intensional como extensional. Como observa González Lagier (2020, p. 92 nota 13), las expresiones propuestas «no logran superar la remisión a factores subjetivos (como la capacidad de imaginar historias o de pensar hipótesis plausibles a favor de la inocencia), no avanzan mucho en la reducción de la vaguedad intensional (remiten a nociones vagas como "si es creíble", "plausible", "difícil de explicar", sin indicar cuándo debe ser creíble, plausible, etc.) y no solucionan el problema de vaguedad gradual (los conceptos a los que se remiten: credibilidad, plausibilidad, etc., son claramente conceptos graduales)». Dei Vecchi (2020b, p. 422) advierte lo mismo en la propuesta de Laudan referido a la «alta dificultad de explicar ciertas pruebas, a lo fiable de ellas, a lo plausible de ciertas historias, a la posibilidad personal de concebir historias alternativas o no, a lo razonable de las hipótesis alternativas (personalmente) imaginables».

[13] Ver más adelante nuestras observaciones sobre la noción de prototipo.

puede diferir en términos diacrónicos) y variables (reconocer que dos sujetos pueden contar con las mismas pruebas, pero llegar a resultados probatorios diferentes). Desde esta perspectiva, lo más adecuado es concebir a los EdP como mecanismos cuyos criterios epistémicos sirven de anclaje para estructurar prácticas argumentativas[14].

Esto nos lleva al segundo punto. Para los efectos de formular un EdP, el recurso al juzgador parece ineludible, por lo que no es cierto que el reconocimiento de la subjetividad haga que las conexiones entre la prueba rendida y los hechos probados pierdan valor. Intuimos que el problema es alimentado por ciertas perspectivas que, por un lado, desconocen la participación del sujeto cognoscente en la formación del conocimiento y, por otro lado, ignoran que nuestras búsquedas están condicionadas por las herramientas que nos permiten justificar nuestros hallazgos (Gutting, 1980, p. 221; Mazzarese, 1997, pp. 587-616).

Laudan parece olvidar que sus propuestas sobre EdP genuinos contienen criterios que no son propiedades intrínsecas de los objetos, sino una característica que se construye entre la evidencia disponible y las habilidades argumentativas del tribunal. Es precisamente en ese espacio de decisión (o de juicio), donde el tribunal debe considerar las particularidades propias del caso concreto[15]. En esta línea, Gama (2021, párr. 80-81) destaca que las formulaciones en que se expresan los EdP son necesariamente subjetivas y epistemológicas, ya que enfatizan (con razón) tanto la calidad de las pruebas como el impacto que generan en quien debe decidir. Dicho autor plantea que pruebas y actitudes subjetivas no son elementos excluyentes y que no es necesario condenar al irracionalismo a teorías que incorporen la convicción del tribunal. En sus palabras, «una teoría que no da cuenta de las creencias o de la convicción del órgano jurisdiccional es una teoría del razonamiento probatorio incompleta».

Pues bien, la búsqueda de la objetividad *tout court* provoca un alejamiento sobre cómo se razona en términos probatorios en los procesos judiciales. La pretensión de Laudan ignora que el razonamiento probatorio no puede *congelarse* en formulaciones abstractas que sacrifican la adaptación del razonamiento a las posibilidades[16]. Por

---

[14] Dei Vecchi, (2022, p. 340), y Gama, (2021, párr. 6), sostienen que los EdP deben ser entendidos como exigencias argumentales que, si bien no garantizan objetividad en el grado de justificación epistémica, permiten identificar los diferentes puntos, temas o aristas que deben ser abordados en las decisiones. La idea es mirar la motivación judicial para diseñar metodologías útiles para valorar las pruebas.

[15] En contra, Aguilera (2023, pp. 256-257) plantea que la restricción de significados solo se justifica bajo una concepción cognoscitivista de la prueba, aunque también existen concepciones que serán «análoga a la creación de una receta culinaria, cuyo contenido depende de las meras preferencias subjetivas de quien la inventa». Sin embargo, no es necesario ir tan lejos y sostener que todo es subjetivismo, sino simplemente destacar que no es necesariamente cierto que exista una interpretación que garantice un vínculo estrecho entre lo pretendido por el legislador al momento de regular un EdP y lo interpretado por el tribunal. En otras palabras, siempre puede haber una disonancia entre la concepción de la prueba asumida por el legislador al momento de regular un EdP y la concepción de la prueba asumida por el tribunal al momento de aplicarlo.

[16] Estimamos que lo que tensiona a Laudan, es algo similar a lo que plantea Carlo Ginzburg (2013, p. 220) acerca de lo que cabe esperar de disciplinas lejanas a su sólida física: «La orientación cuantitativa

eso, si bien la crítica a los EdP formulados en términos exclusivamente subjetivos es pertinente, una idea de total objetivación carece de asidero: la aplicación de los EdP está inevitablemente mediada por la interpretación y los juicios valorativos de los operadores jurídicos. De esta manera, es preciso reconocer que la argumentación probatoria también depende de factores circunstanciales, como las destrezas y capacidades de quienes toman las decisiones, o de las creencias compartidas de la comunidad dentro de la cual se opera[17]. Esta peculiaridad evidencia que los hechos probados, además de derivar del análisis de la prueba disponible, son el reflejo de las formas en que los criterios epistémicos son regulados e interpretados[18].

Un breve comentario más para cerrar este punto. Recientemente Coloma y otros (2024, párr. 48) han propuesto concebir al razonamiento probatorio desde una *perspectiva situada*. A nuestro entender, el carácter situado está mediado por la manera en que los individuos conciben ciertas actitudes o situaciones y, en definitiva, por aspectos determinados por la experiencia, género, clase social, entre otros. En ese sentido, el acto de conocer y de establecer la suficiencia probatoria no sería una actividad neutra y desvinculada de los sujetos, sino ligada al contexto social donde se realiza tal actividad. Asumir que la realidad puede ser observada de forma neutra es también asumir que todas aquellas personas que no cumplen con un «estándar de normalidad» (o ciertas credenciales) se encuentran ante una dificultad para situarse en la calidad de observadores honestos[19]. Por eso, perspectivas que se inscriben bajo una concepción universal de sujeto cognoscente, donde el tribunal debería operar aislado de su ambiente social y desatendiendo de su contexto histórico, no representan la realidad de la toma de decisiones probatorias. Así las cosas, el acceso al pasado está mediado por el punto de vista (*standpoint*) de quien conoce, lo que implica tomar partido por una visión del mundo, con distintos compromisos sobre lo que es (y debe ser) considerado como verdadero. En este marco, Gama (2020, p. 290) advierte que la inserción del sujeto cognoscente en su entorno «mantiene abierta la crítica ahí donde la concepción racional de la prueba parece insistir en la pretensión de neutralidad y objetividad de los métodos y presupuestos»[20].

---

y antropocéntrica de las ciencias de la naturaleza, desde Galileo en adelante, ha llevado a las ciencias humanas ante un desagradable dilema: o asumen un estatus científico débil, para llegar a resultados relevantes, o asumen un estatus científico fuerte, para llegar a resultados de escasa relevancia».

[17] Coloma y otros, (2021, p. 136). Azócar (2018, p. 24), por su parte, ha estudiado cómo las emociones de jueces y juezas de familia impactan en la valoración de testimonios y en la acreditación de hechos.

[18] Gascón (2010, p. 110) recuerda que, gracias a la teoría epistemológica postpositivista, las normas procesales y probatorias pueden entenderse «como los esquemas teóricos desde los que se conocen los hechos en el proceso».

[19] La idea de «estándar de normalidad» se asocia a una concepción de estándar como prototipo desarrollada más adelante.

[20] Tuzet (2021, p. 80) recuerda que la justificación probatoria se elabora dentro de una práctica social, donde no cuenta la opinión individual, sino lo que la comunidad reconoce como justificación (ciertas fuentes, ciertos métodos, etcétera).

Como a continuación veremos, la dispersión de resultados probatorios no puede ser suprimida, sino sólo atenuada mediante ciertas *medidas de higiene*[21] que toman la forma de procedimientos y criterios más estructurados. Por esto, aunque es posible buscar una mayor estandarización con nuevas formulaciones de EdP, no representa un avance sostener que la determinación de la suficiencia probatoria deba desligarse de los juicios de quien debe decidir[22].

3. LOS EDP NO SIRVEN PARA DISTRIBUIR FALSOS POSITIVOS Y FALSOS NEGATIVOS

Laudan (2013, p. 23) sostiene que «un sistema de justicia penal es primordialmente un *motor epistémico*, es decir, un dispositivo o herramienta para descubrir la verdad a partir de lo que a menudo comienza con una mezcla confusa de pistas e indicios». Consciente de la dificultad de alcanzar ese objetivo, defiende que las normas probatorias deben diseñarse no sólo para minimizar el error, sino también para distribuirlo entre las partes[23]. Esto se debe a que no es posible «eliminar la incertidumbre por completo, ni hacer que las inferencias abductivas que se emplean en el razonamiento jurídico preserven siempre la verdad» (2013, p. 130)[24]. Así, Laudan propone dejar de lado «las preguntas relacionadas con el papel que los valores no-epistémicos desempeñan en la administración de justicia» (2013, p. 26), para adherirse «al punto de vista de que cuanto menos se hable de derechos, tradiciones y doctrinas jurídicas y derecho constitucional, será mejor» (2013, p. 29).

Para Laudan, el error se produce cuando hay discrepancia entre lo que se ha probado dentro del proceso y lo que realmente ocurrió. Estos errores conforman los llamados «veredictos falsos» que, en el proceso penal, se manifiestan de dos formas: «cuando una persona inocente es tratada como culpable [falso hallazgo inculpatorio o también falsos positivos] o cuando una persona culpable no es hallada como

---

[21] Kahneman, Sibony y Sunstein,(2021, pp. 271 y ss).

[22] Dei Vecchi (2023, párr. 52) advierte que confiar en la posibilidad de fijar legislativamente umbrales de suficiencia probatoria capaces de eliminar el juicio valorativo judicial es, en definitiva, un *wishful thinking*.

[23] También asume una mirada naturalista según la cual «la mayoría de los asuntos de interés filosófico, en última instancia, descansan en averiguar cuáles son los hechos relevantes y dependen de ello» (Laudan, 2013, p. 31). A pesar de ello, considera que, además del EdP, la presunción de inocencia, el beneficio de la duda y la carga de la prueba depositada en la acusación conforman la *doctrina de la distribución del error*. Así, aunque ninguno de aquellos conceptos «está inspirado en la búsqueda de la verdad deben, no obstante, ocupar un lugar central en un libro de epistemología del derecho procesal penal» (Laudan, 2013, p. 60).

[24] Recientemente Dei Vecchi (2023, párr. 39) ha destacado que la verdad de la premisa fáctica no es condición necesaria de la justificación externa de dicha premisa, aun cuando esa justificación apunta a alcanzar dicha verdad. Como la justificación probatoria es parte del género de la justificación epistémica, «al igual que sucede en cualquier otro ámbito de la investigación empírica, un enunciado acerca de ciertos hechos puede estar justificado siendo, a pesar de ello, falso».

tal por el sistema [falso hallazgo exculpatorio o también falsos negativos]» (2013, pp. 33-34). En este contexto, el efecto esperado del EdP es distribuir los errores de acuerdo con las inclinaciones político-morales de la sociedad e, incluso, reflexionar sobre la asignación de un riesgo de error simétrico (o no) a las partes del proceso. De este modo, la formulación de un EdP requiere, por un lado, determinar cuál es la distribución del error deseada desde una perspectiva político-moral y, por otro lado, establecer un umbral de suficiencia que materialice dicha distribución (Dei Vecchi, 2020, p. 421).

Como se sabe, salvo que se asuma una concepción cognoscitivista de la moral, el reparto adecuado del riesgo y el umbral de suficiencia son asuntos definidos de forma contingente. Determinar el *quantum* de suficiencia probatoria es una decisión que difícilmente tiene una respuesta unívoca, ya que los bienes jurídicos en juego gozan de una importancia social diferente (Ferrer, 2025, pp. 140-152). Así, conforme a Laudan (2013, p. 105), «el umbral se establece más allá del punto medio sólo cuando creemos que una de las clases de error que pueden cometerse es más costosa —por lo que nos interesa que se evite con mayor asiduidad— que la otra». Esto es de particular importancia para el funcionamiento del sistema procesal penal ya que, si la formulación de un EdP para condenar es oscura, no contribuye mucho a inspirar confianza en el sistema (Laudan, 2013, p. 88)[25]. De ahí que los EdP conecten la epistemología con la moralidad: la primera proporciona los medios que consideramos adecuados para conocer y la segunda proporciona el umbral que consideramos adecuado para tomar una decisión[26].

En resumen, un genuino EdP debiese operar como un mecanismo adecuado para distribuir el error entre las partes de un proceso mediante el establecimiento de un umbral de justificación adecuado para adoptar una decisión por sobre otra, asociando un costo a cada tipo de error frente a posiciones dicotómicas, como la de condenar o absolver. No dudamos que esta tesis puede ser útil para explicar una faceta de la discusión relativa a las finalidades del proceso penal, pero su alcance es bastante más restringido al momento en que los tribunales deben aplicar los EdP.

Nuestras dudas se originan en el concepto de error que, a juicio de Laudan, sería posible «distribuir». La idea según la cual podemos distribuir falsos positivos y negativos mediante un EdP parece suponer que estamos en condiciones de identificar, en

---

[25] Conforme a Ferrer (2025, pp. 34-36) no se trata un umbral probabilístico o cuantificado, sino de uno que reduzca en todo lo posible la vaguedad que todo criterio no matemático tendrá. La idea es lograr una mayor cantidad de casos claros de (in)cumplimiento del estándar y un menor margen de casos dudosos.

[26] Céspedes (2021, p. 253) argumenta que el EdP no distribuye los errores dentro del proceso, sino (i) materializa la importancia que le asignamos a la posibilidad de causarle daño moral a un acusado al condenarlo erróneamente (derivada del deber de no dañar injustamente a otro) y (ii) otorga un tratamiento igualitario a los miembros de la comunidad. Un criterio cualitativo «no responde a ninguna relación entre falsos positivos y falsos negativos, sin perjuicio de que, contingentemente, de la utilización del criterio adoptado resultará —al menos desde un punto de vista estadístico— una determinada relación entre dichos extremos».

alguna medida, cuándo acontece un veredicto falso, es decir, cuándo se da el caso de que lo afirmado formalmente en el proceso no coincide con aquello que ha ocurrido en la realidad. Así, las objeciones que nos interesa destacar son las siguientes: (1) no es posible distribuir algo que desconocemos y (2) es posible llamar error judicial no solo a decisiones que (suponemos) se fundamentan en premisas falsas, sino también a conjuntos de decisiones dispersas sobre asuntos similares carentes de un fundamento aceptable para explicar tal dispersión[27].

Con respecto al primer punto, es preciso distinguir dos dimensiones de los EdP: aquella que los entiende como productos de una decisión de política procesal, y aquella que los concibe como guía para la acción de los tribunales en los casos que resuelven. Por un lado, los EdP se presentan como una decisión política sobre la cantidad de errores (por ejemplo, condena de inocentes o absolución de culpables) que la sociedad está dispuesta a tolerar. Es una determinación *a priori* sobre la tasa de distribución de errores en los múltiples juicios que serán realizados. Así, al momento de evaluar si la aplicación del MADR ha sido exitosa, debemos estar en condiciones de afirmar, por ejemplo, que se ha condenado solo un 4 por 100 de inocentes; y que, en cambio, el 20 por 100 de culpables han sido absueltos. El éxito radicaría en que la disposición a absolver a un número considerable de culpables logra mantener el porcentaje de condenas erróneas de inocentes por debajo del 4 por 100. Por otro lado, la situación será distinta al analizar el cumplimiento del MADR en un caso concreto. Aquí la preocupación es determinar la probabilidad de condenar a un inocente antes de dictar sentencia. Si el porcentaje es superior al 4 por 100 debemos absolver; si es inferior, debemos condenar. Así, si al decidir, el porcentaje de error de condena es de 3 por 100 y la persona en realidad era inocente, aunque la condena fue errónea desde la perspectiva de la verdad, no lo fue desde la perspectiva de la aplicación del MADR (estábamos en una posición que justifica la decisión de condenar).

A nuestro juicio, la pretensión del párrafo anterior no tiene sentido dado el actual estado de las cosas: no sabemos el porcentaje real de inocentes condenados ni de culpables absueltos. El estado de ignorancia sobre el error desde el punto de vista de la verdad es un dato inaccesible[28]. Para profundizar en ello distinguiremos entre un punto de vista interno y externo al proceso judicial. El primero se relaciona con el

---

[27] Piense, por ejemplo, en las decisiones donde un tribunal dispone que no procede la prisión preventiva solo para que, días más tarde, un tribunal superior establezca, en la misma causa, la necesidad de privar de libertad al imputado mientras dure la investigación. Aquí el error no parece ser un sinónimo de falsedad, sino de dispersión injustificada. Tenemos dos afirmaciones relativas al mismo caso que parecen incompatibles (no con la realidad, sino) entre sí.

[28] Para Ferrer (2025, pp. 134-138), como no es posible contar con información empírica sobre «la cantidad y distribución de verdaderos inocentes y verdaderos culpables que son enjuiciados», no es posible identificar «mecanismos causales directos que vinculen los estándares de prueba con ratios específicas de distribución de decisiones erróneas». Tampoco es posible tener «información empírica fiable sobre esas mismas ratios, de modo que resulta del todo inviable la estrategia de fundar el nivel de exigencia probatoria de los estándares de prueba en las ratios de distribución de los errores a las que supuestamente darían lugar».

punto de vista de quienes participan en un proceso buscando cumplir con una serie de objetivos institucionales. El segundo, en cambio, con el punto de vista de quienes, interesados por la correcta aplicación del derecho, observan el funcionamiento del sistema judicial sin participar en la dinámica procesal. Esta distinción es importante porque nos permite comprender el sentido o alcance que podemos adscribir a la tesis según la cual los EdP son mecanismos adecuados para distribuir el error. En particular, la distinción implica que el concepto mismo de error utilizado en tal tesis no puede significar lo mismo desde un punto de vista interno que desde un punto de vista externo.

La distribución de errores, asumiendo un punto de vista interno del proceso, parece difícil de sostener porque la aproximación a la realidad en instancias judiciales se produce, precisamente, a partir de la prueba disponible. Luego, afirmar que existe una discrepancia entre lo probado y lo verdadero, sugiere que estamos en condiciones de propiciar una aproximación a la realidad diferente a la que se produce mediante la prueba. Así, la idea de que los hechos verdaderos (lo que «realmente ocurrió») pueden operar como un criterio útil para evaluar la corrección de las decisiones sobre hechos probados no es viable porque, dentro del proceso judicial, no tenemos acceso a los primeros como algo diferente a los segundos.

Lo anterior conduce a que, al interior del proceso judicial, decimos que el juez ha cometido un error si es posible realizar acciones que impiden que la sentencia adquiera el estatus de cosa juzgada. Así, es posible interponer recursos para mantener abierta la discusión (como la apelación o la casación) para sostener que el tribunal de instancia ha errado. En estos casos, error no significa falsedad, ya que los sistemas recursivos (salvo circunstancias excepcionales como la acción de revisión que recae sobre una sentencia que ya produce cosa juzgada) no exigen estar en condiciones de sostener que ha existido una discrepancia entre lo que ha afirmado el juez y lo que ha ocurrido realmente. Lo que se exige es algo mucho más modesto: sostener que la decisión no se encuentra justificada en relación con las reglas que regulan el procedimiento probatorio. Entonces, desde un punto de vista interno, un concepto de error como sinónimo de falsedad establece un baremo inoperante para cumplir con las funciones que asignamos a los EdP dentro del proceso judicial.

Se podría intentar defender a Laudan argumentando que parece estar más interesado en el punto de vista externo al proceso judicial (recordemos que, a su juicio, el proceso penal podía ser *una parodia de sistema de prueba*). Es decir, asume que la discrepancia entre lo probado y lo verdadero puede ser identificada comparando el resultado formal del proceso con otro resultado obtenido mediante una aproximación a la realidad que no se encuentra condicionada por reglas procesales subóptimas. De esta forma, en un contexto sin las reglas que favorecen la aparición de errores en el proceso penal, podríamos llegar a resultados que nos permitan contrastar lo probado y lo «real»; siendo lo «real», en este caso, aquello que ha sido probado en un escenario hipotético con reglas epistémicas adecuadas. Si este es el caso, la tesis de Laudan parece resistir, pero con un alcance limitado. Resulta curioso, en este sentido, que la principal función de un EdP no tendría relación con los mecanismos procesales

que permiten el control de las decisiones probatorias; y no resulta particularmente útil para comprender el error cuando este se refiere, no a la realidad, sino que a la suficiencia de la prueba disponible. Si «error» es sinónimo de «falsedad», entonces, la distribución de errores sería un dato accesible solo para quienes están fuera del proceso; e irrelevante para quienes están dentro[29].

Nuestra segunda objeción apunta a la falta de reconocimiento de errores entendidos como conjuntos de decisiones dispersas sobre asuntos similares en ausencia de un fundamento aceptable para explicar tal dispersión. Para abordar esto, recurriremos a las ideas de Kahneman, Sibony y Sunstein (2021, p. 37) en torno a la relación que existe entre las nociones de «juicio» y «error». Para los autores, «las cuestiones de juicio, incluidas las profesionales, ocupan un espacio intermedio entre las cuestiones de hecho o de cálculo, por un lado, y las cuestiones de gusto u opinión, por otro». Mientras que las cuestiones de cálculo admiten una única respuesta correcta, las cuestiones de juicio están definidas por la expectativa de un desacuerdo limitado, donde el «grado exacto de desacuerdo aceptable es en sí mismo un juicio subjetivo que depende de la dificultad del problema»[30]. Así, las prácticas disciplinares suelen exigir que los distintos juicios de sus participantes estén en el marco de las respuestas posibles y que las diferencias puedan ser explicadas. Cuando ello no ocurre, la práctica pierde credibilidad en tanto los juicios de sus participantes se parecen demasiado a simples opiniones.

Es posible que al momento de pensar en el correcto funcionamiento del sistema judicial, Laudan tuviese en mente que las decisiones probatorias pueden representarse, de manera exitosa, como cuestiones en donde solo existe una única respuesta correcta para cada caso. Su noción de error da buena cuenta de ello. Como se sabe, el sistema penal debe adoptar una decisión entre dos posiciones opuestas, condenar o absolver. Las posibles decisiones presuponen, a su vez, una justificación estructurada en términos dicotómicos: el imputado cometió o no cometió el delito que se le imputa; luego, la decisión de condenar o absolver puede (o no) ser correcta si se corresponde (o no) con la realidad. En este esquema de pensamiento, toda dispersión es siempre indicativa de que se ha cometido un error. No es posible que dos jueces,

---

[29] Es preciso aclarar que, incluso, si aceptamos que a Laudan le interesa solo la faceta externa de la discusión probatoria, podrían subsistir legítimas dudas de que error sea sinónimo de falsedad. Para que aquello fuese operativo debiese debilitarse su posición sobre lo que es un falso positivo y un falso negativo, en cuanto a que, tendría que ser algo contingente al sistema probatorio externo al que se recurre y no un asunto de correspondencia con el mundo.

[30] Para los autores, las cuestiones de «hecho o de cálculo» se definen en relación con la existencia de un valor que resulta posible verificar en el mundo para los efectos de establecer si ha existido o no un error. Por ejemplo, la afirmación de que «las dolencias del paciente se deben a un desgarro en la pierna derecha» puede verificarse si consultamos las radiografías que acabamos de tomar al paciente. Luego, es el conjunto de radiografías las que operan como criterio de la existencia de un error. Este tipo de evaluación puede darse ocasionalmente y, en general, sobre un punto específico de prueba de los procesos judiciales. Sería el caso de «X es el padre de Y, pues así se confirma del examen de ADN». Al ser tan bajo el margen de error, no hay gestión de la incertidumbre, ni tampoco dispersión en las decisiones: se trata de un escenario en que los EdP carecen de relevancia.

aplicando el mismo EdP al mismo caso, tomen decisiones diferentes sin que la respuesta de uno de ellos sea incorrecta. Incluso si no sabemos cuál es el juez que se ha equivocado, podemos estar seguros de que ha tenido lugar un error.

En conformidad a lo señalado, resulta más satisfactorio discutir sobre la suficiencia probatoria en los términos de quien decide bajo la expectativa de que su decisión no se aleje demasiado de las decisiones habitualmente toleradas por el sistema. Aquí, el error no implica falsedad, sino dispersión injustificada: un juicio es errado cuando se aleja de las pautas de decisión aceptadas por los miembros de una comunidad disciplinar[31]. Esto nos obliga a conformarnos con analizar la fundamentación de la decisión adoptada y a reconocer que es perfectamente posible que exista más de una respuesta correcta para las cuestiones fácticas debatidas en juicio.

Por cierto, resulta extraño hablar de la existencia de dos o más respuestas correctas en el terreno probatorio (salvo que las diferencias entre ellas sean de matices). Para entender mejor esto, aclaramos que entendemos que una decisión es correcta cuando es compatible con la normativa aplicable (EdP incluido). La respuesta correcta sería, entonces, aquella que ha respetado las exigencias del sistema y no es susceptible de anular. Esto se puede ver con claridad si observamos lo que sucede en otros ámbitos, como la historiografía, donde no se trata de contraponer «lo verdadero a lo inventado sino la integración, puntualmente señalada en toda ocasión, de "realidades" y "posibilidades"» (Ginzburg, 2010, p. 439). Es cierto que, en el caso de los procesos judiciales, no pareciera haber tanta flexibilidad para arriesgar diferentes relatos que cuenten con credenciales para reconstruir lo posible a partir de la prueba rendida, pero aquello no está excluido. Son precisamente los EdP quienes determinan hasta dónde podemos llegar.

Pues bien, a pesar de que los desacuerdos en la aplicación de los EdP son indeseables, a nuestro entender también son inevitables: distintas personas pueden llegar a decisiones divergentes bajo circunstancias idénticas. Por esto, si bien es adecuado instaurar *medidas de higiene* (como las propuestas por el propio Laudan o Ferrer) es importante entender que no garantizan respuestas que resistan cualquier objeción epistémicamente fundada, pero sí contribuyen a estandarizar las decisiones. De ahí que las propuestas puedan servir como mecanismos para que las decisiones probatorias no dependan *en exceso* de la presunta experticia o intuición de los operadores jurídicos, sino que promuevan un enfoque basado en el seguimiento de procedimientos estructurados. En otras palabras, aunque no somos tan optimistas a los efectos de neutralizar toda potencial libertad en la argumentación sobre la suficiencia probatoria, sí reconocemos que es posible compactar la variabilidad de las decisiones (el ruido)[32].

---

[31] Véase la noción de prototipo más adelante.

[32] En ese sentido, al momento de establecer una nueva legislación, la autoridad normativa se pregunta acerca de qué formulación se presenta a menores manipulaciones en sede interpretativa. Esto implica reconocer que la interpretación depende de los presupuestos teóricos con los que cada intérprete se aproxima a las disposiciones normativas. Guastini, (1999, p. 34).

Antes de cerrar, vale la pena aclarar que no objetamos la idea muy gruesa de que los EdP inciden en los falsos positivos y falsos negativos que arroja el sistema judicial en su conjunto. En tal sentido, un EdP muy exigente debiese arrojar menos falsos negativos y más falsos positivos que uno poco exigente. Así, lo que ponemos en duda es que, en el terreno de la práctica, los EdP tengan la función primordial de distribuirlos. No olvidemos que lo que se debate en juicio es la elección de uno entre dos o más enunciados representacionales, y son las pruebas disponibles las que sirven para decidir. Tal elección implica la generación de hechos probados que resultan de una serie de operaciones institucionalizadas que adquieren sentido, precisamente, porque desconocemos lo que realmente ocurrió. Por esto, en rigor, los hechos probados no son una reproducción de una realidad externa, sino conclusiones epistémicas elaboradas mediante procedimientos institucionales y normas procesales[33]. El punto interesante es que más allá de una serie de funciones indirectas de los EdP (cuánta prueba hay que producir, de qué magnitud pueden ser los saltos inferenciales en los discursos de justificación, etc.)[34], ellos son importantes para reducir el ruido en los sistemas de adjudicación, acorde a las directivas del legislador, conforme a un juicio de semejanzas y diferencias. Sobre esto seguiremos profundizando en lo que sigue.

4. LOS EDP NO (SOLO) SON RECONSTRUIBLES COMO UMBRALES

En consonancia con lo ya indicado, Laudan sostiene que los EdP deben definir el grado de confirmación requerido para que una hipótesis se tenga por probada. Para ello concibe a los EdP como un mecanismo que estipula «cuándo la relación que existe entre (…) las premisas y la conclusión que se busca justifica la aceptación de dicha conclusión como probada en el contexto específico» (2013, p. 124)[35]. Así,

---

[33] Aunque Laudan (2013, p. 35) rechaza que los fallos crean la culpabilidad o inocencia de las personas, consideramos que es más preciso entender el hecho probado como una construcción institucional que surge al superar las exigencias procesales. En ese sentido, lo determinante para la resolución del conflicto es aquello que queda establecido dentro del proceso tras respetar los procedimientos, cerrar definitivamente la discusión y bloquear nuevas problematizaciones. Concebir al hecho probado como una construcción no conduce a un relativismo puesto que, lejos de ser individuales, las construcciones son histórico-sociales. Conforme a Tuzet (2021, pp. 92-93), construcción no significa creación, sino elaboración a la luz de categorías jurídicas, o selección de elementos fácticos jurídicamente relevantes. En su modelo cognoscitivista, Gascón (2010, p. 108) reconoce que en el ámbito judicial la cuestión sobre la verdad debe ser cerrada autoritativamente. En sus palabras: «claro está que esta verdad última no será infalible, pero será final, en el sentido de que pondrá fin al conflicto autorizadamente. Por eso, según vimos, alguna razón tenía Kelsen al hablar de la naturaleza constitutiva de la sentencia».

[34] Un desarrollo amplio de lo indicado puede verse en Coloma, 2016a, pp. 36-47.

[35] En el mismo sentido, Gascón (2005, p. 134): «en lugar de indicar numéricamente el grado de probabilidad que debe alcanzar una hipótesis sobre un hecho, [los EdP] deberán indicarnos cuándo está justificado aceptar dicha H como verdadera; es decir, cuáles son los criterios que establecen cuándo se ha conseguido la prueba de un hecho».

los genuinos EdP, en el sentido en que Laudan los entiende, debieran indicar con claridad un «punto de quiebre» que distinga los enunciados que han sido probados de aquellos que no lo han sido. Para ello recurre a la metáfora[36] de los *umbrales*[37].

¿Es esta la mejor manera de representar los EdP? ¿Captura la metodología de toma de decisiones en situaciones de incertidumbre? Nuestra objeción aparece al momento de contrastar la propuesta de Laudan con la forma en que habitualmente procesamos las preguntas relativas a la vaguedad extensional de una categoría. Aquello tiene sentido porque los EdP nos sirven para determinar si un cierto ejemplar dudoso debe o no ser incorporado dentro de una categoría relevante («los que han matado a otro», «los que han provocado un daño ambiental», etc.). De esta forma, si se nos pide, por ejemplo, determinar si «X» cuenta como un «ave», o si «Y» cuenta como un «buen libro», parecen existir dos caminos posibles; y solo uno de ellos coincide con la metáfora de los umbrales.

El primer camino consiste en listar un número de propiedades necesarias para incluir un caso dentro de una categoría. Si tenemos dudas sobre la inclusión de «X» en la categoría de «ave» (o testigo creíble), debiésemos corroborar que las propiedades definitorias del concepto de ave (o testigo creíble) estén presentes en «X». Por cierto, algunas de las categorías que nos interesan podrán padecer del problema de la vaguedad combinatoria, esto es, podría no ser necesario que concurran todas las propiedades que en principio se vinculan a ellas (no todas las aves vuelan, o no todos los testigos creíbles son capaces de mantener un relato coherente). Otra opción es que baste que una o más propiedades se den *hasta cierto punto*. Esto último supone transformar en porcentajes el nivel de concurrencia requerido de cada una de las propiedades definitorias, lo que no excluye el uso de conceptos graduables en lugar de números (Clermont, 2013, pp. 35 y ss.). En otras palabras, este camino supone operar con algo equivalente a un *check list* y, si se cumple con los valores prefijados, se entenderá que el caso originariamente dudoso entra en la categoría que interesa (en nuestro caso, la de hecho probado).

Aun cuando esa forma de operar pareciera asegurarnos un método fiable para tomar decisiones, algunos filósofos de la mente y científicos sostienen que no es la manera habitual en que nuestro cerebro categoriza fenómenos. La psicóloga y neurocientífica Lisa Feldman Barrett (2018, p. 152) sostiene lo siguiente:

> De las cenizas de los conceptos clásicos nació una nueva visión que defendía que un concepto se representa en el cerebro como el mejor ejemplo de su categoría, al que se llamó «prototipo». Por ejemplo, el ave prototípica tiene plumas y alas y puede volar. No todos los casos de «Ave» tienen

---

[36] En cuanto al contenido metafórico del uso del concepto de umbral, véase Agüero, (2016, pp. 83 y ss.).

[37] Laudan no aborda de lleno el problema que aquí se plantea en cuanto a que para hablar EdP, la manera adecuada de hacerlo es por la vía de recurrir a la metáfora de los umbrales. Lo que ocurre es, simplemente, que Laudan parece darlo por descontado. Al focalizar el problema de los EdP en la suficiencia probatoria, no da cuenta de una aproximación diferente como la que aquí se realiza desde la metáfora de los prototipos. Para la relación entre EdP y umbrales, véase Páez, (2015, pp. 129 y ss.).

> estas características, como los avestruces y los emús, pero siguen siendo aves. La variación en relación con el prototipo es posible, pero siempre que no sea demasiada: una abeja no es un ave aunque tenga alas y pueda volar. Según esta visión, cuando aprendemos una categoría se supone que nuestro cerebro representa el concepto como un solo prototipo. Podría ser el ejemplo más frecuente de la categoría o el ejemplo más típico, es decir, el caso que encaja más con la categoría o que tiene una mayoría de sus características.

Así, la idea de que los estándares pueden representarse solo a modo de umbrales resulta contraria a nuestra experiencia, pues la regla general es que no estamos en condiciones de solucionar los problemas extensionales mediante un ejercicio de medición. A nuestro juicio, no toda pregunta puede responderse exitosamente mediante un análisis cuantitativo porque no todo fenómeno puede ser medido. Además, la medición es un acto costoso en términos epistémicos: requerimos de una escala y de suficiente claridad sobre las propiedades que estamos midiendo en cada caso. De esta forma, la representación de los EdP a modo de umbrales parece cubrir solo parcialmente la manera en que normalmente resolvemos casos dudosos sobre la extensión de una categoría.

> Agúero (2016, pp. 85-86) ilustra el problema del uso de umbrales aludiendo a la historia de la medición de la temperatura. Antes de disponer de escalas de medición como las que hoy utilizamos y que miden la energía térmica mediante instrumentos calibrados entre sí, lo que se hacía era simplemente considerar la sensación térmica. Y ello, por cierto, no es lo mismo.
>
> Sigamos con ese ejemplo: para la toma de decisiones a partir de la temperatura entendida como energía térmica la idea de umbral funciona perfecto (*déle un analgésico a la niña cuando la temperatura le aumente a 38°*). En cambio, para la toma de decisiones a partir de la temperatura entendida como sensación térmica la idea de umbral deja de funcionar (al menos en términos fuertes, pues no contamos con °Celsius, ni Fahrenheit) y es necesario recurrir a algunas señales específicas (déle un analgésico cuando le den escalofríos) o a un prototipo (*déle un analgésico a la niña cuando sienta su frente parecida a la de una persona que lleva dos horas cocinando en pleno verano y sin una brizna de viento*) [38].

El atractivo que posee el uso de la metáfora de los umbrales, tal vez, radica en su consistencia con la tendencia hacia la fragmentación de nuestros objetos de análisis. En el caso de Laudan (2005, pp. 96-97) aquello queda bien reflejado cuando en el inicio del artículo al que ampliamente hemos referido distingue entre distintas clases de problemas teniendo como patrón de diferenciación su relación con lo epistémico. Al considerar que los EdP quedan dentro de aquellos conceptos que nos sirven para proteger los valores que califica como quasi epistémicos, pareciera verse compelido a diferenciarlo de todo aquello que sirve para proteger los valores epistémicos (especialmente regulación de la valoración de la prueba). Esto, a nuestro juicio, es un problema pues fuerza a fijar momentos nítidamente diferenciados para aludir a actividades que *en la práctica* son difíciles de diferenciar (valorar la prueba y aplicar

[38] Lo expresado no parece absurdo. Así, antes de encontrar puntos de referencia para medir la temperatura se propusieron prototipos tales como *el frescor de un sótano parisino*. Seth, (2023, p. 50).

el EdP). Aquello no es inofensivo pues nos plantea el difícil problema de establecer unidades de medición separadas de lo que sería los criterios del análisis[39].

Para reforzar lo anterior, permítasenos recurrir a un breve fragmento del ensayo de Thomas Kuhn, *La función de la medición en la física moderna* (1993, pp. 217 y 237). Allí se dice que: «Es frecuente que el científico no pueda obtener cifras que concuerden con la teoría mientras no sepa qué cifras debe hacer que produzca la naturaleza» y más adelante se añade:

> «[…] normalmente, es condición previa, para una cuantificación fecunda de un campo de investigación dado, una gran cantidad de investigación cualitativa, tanto empírica como teórica. Sin tal trabajo previo, la directriz metodológica "Salgamos a medir" puede resultar tan solo una invitación a perder el tiempo».

El análisis cuantitativo es atractivo, pues cuando está disponible permite un nivel de control elevado sobre la discrecionalidad/arbitrariedad del tribunal. El problema es que la medición propia de un umbral solo en unos pocos casos judiciales está disponible.

El segundo camino para zanjar el problema de si un ejemplar dudoso pertenece a una categoría consiste en emplear un prototipo para realizar un análisis cualitativo de semejanzas y diferencias. Ello ocurre cuando se formula, de manera hipotética, un caso que indubitadamente pertenece a una categoría. Imaginemos entonces que Ud. desea determinar si un libro cualquiera, por ejemplo, alguno perteneciente a la saga de *Harry Potter*, pertenece o no a la categoría de «buenos libros». Dado que no existe un criterio cuantitativo útil para este ejercicio (no vale contar el número de páginas o la cantidad de copias vendidas), usted debiese recurrir a una comparación a partir de un ejemplo indubitado. Según su percepción, nadie duda de que *Los Hermanos Karamazov* es un ejemplo de lo que cuenta como un buen libro, por lo que nos valdría bien para comparar. Ello exige un esfuerzo argumentativo según el cual se asemeje o distinga un caso determinado con el prototipo que hemos seleccionado; es decir, determinar si *Harry Potter* «se parece» o no a *Los Hermanos Karamazov*, al menos en aquellas características o propiedades que son consideradas como las más relevantes. Dadas las diferencias claras entre ambas novelas, es probable que se requiera hacer algunos ajustes de nuestro ejemplar básico indubitado. Así, podría ser útil buscar algunos ejemplares de libro con características más cercanas al que nos interesa, incluso si la propiedad de ser un buen libro no sea tan indubitada. En dicha dirección podríamos arriesgar de ajustar nuestro prototipo recurriendo, por ejemplo, a *El Señor de los Anillos*.

---

[39] David Bohm (2018, p. 20), un destacado físico teórico, expresa preocupación por la excesiva tendencia a la fragmentación en distintos ámbitos de la vida (la ciencia, por supuesto). Dice textualmente: «La idea que todos estos fragmentos existen por separado es, evidentemente una ilusión, y esta ilusión no puede hacer otra cosa que llevarnos a un conflicto y una confusión sin fin [...] Porque, como ya sabemos, este modo de vivir nos ha abocado a la contaminación, a la destrucción del equilibrio de la naturaleza, a la superpoblación [...]».

Como se puede observar, un prototipo aspira a ser lo suficientemente situado como para arriesgar algunos ajustes al problema concreto que nos preocupa. Para explicar esto recurriremos nuevamente a Feldman Barrett (2018, p. 154). Ella afirma que:

> [L]os prototipos podrían parecer buenos modelos […] de no ser por un detalle paradójico. Cuando medimos casos reales […] este prototipo rara vez se encuentra en la vida real. [...] Si no hay prototipos almacenados en el cerebro, ¿cómo es que la gente enumera sus características con tanta facilidad? Lo más probable es que el cerebro construya prototipos en el acto cuando los necesitamos.

En definitiva, los prototipos con los que operamos no son estáticos, sino que se ajustan al contexto en que requerimos operar. Así, para la determinación de si Fulano es un *buen padre de familia* (o una *persona razonable*), no nos desentendemos, en la medida de lo posible, de su participación en cierta cultura, de su nivel de formación, de sus creencias religiosas, de su orientación sexual, etc. [40]. Conforme a esto, Coloma (2016b, p. 27) ha sostenido:

> El modus operandi de los estándares entendidos en este sentido [como prototipo] se diferencia del [de los umbrales] al tener que realizar el agente un análisis de semejanzas y/o diferencias de los casos dudosos respecto del prototipo. La semejanza implica pertenencia y la diferencia, no pertenencia. El análisis que se lleva a cabo cuando se entiende un estándar como prototipo es cualitativo.

El asunto de si los estándares de prueba quedan mejor representados si se les entiende solo como umbrales, como umbrales o prototipos, o únicamente como estos últimos, debiera decidirse observando prácticas afortunadas de su aplicación. De lo que se trata, entonces, es de identificar uno o más casos en los cuales entendamos estar autorizados a decir que se operó con un EdP, y que esa forma de operar calza mejor con la metáfora del estándar como umbral o como prototipo. A falta de este ejercicio, lo que podemos afirmar es que la metáfora propuesta por Laudan explica solo parcialmente algunas de las expectativas que tenemos al momento de operar con un estándar. Probablemente, puede funcionar bien cuando contamos con suficientes elementos para analizar el caso desde una aproximación que podríamos llamar atomista (las distintas partes nos permiten construir una imagen del todo). Pero no hemos encontrado razones para creer que estas expectativas excluyan otras representaciones, sobre la manera en que se resuelven los casos ubicados en la zona penumbra (donde la mejor aproximación es holista, al ser necesario ver el problema desde la totalidad, ya sea porque los fragmentos se nos escapan, o porque nos hacen perder la complejidad).

Cabe destacar que la apertura a concebir a los EdP como prototipos no implica presentar a los EdP del «modo en que nos venga en gana» (Aguilera, 2023, p. 238). La crítica no debe tomarse a la ligera toda vez que advierte sobre el peligro de la arbi-

[40] Ferraris (2023, pp. 93-94) sostiene que el reconocimiento de ejemplares entraña la realización de un «juicio reflexionante» (o de generalización) sobre características compartidas entre distintos individuos.

trariedad en la aplicación de los EdP, tal como ocurriría si se asumiera que el prototipo se agota en la intuición del decisor que no requiere ser justificada. En este sentido, la representación prototípica de los EdP no es una licencia para la discrecionalidad ilimitada, sino una herramienta que permite una mejor adaptación de los estándares a la realidad de la práctica judicial. Siempre se requiere de una argumentación que explique en qué medida el caso en cuestión se asemeja o se diferencia de los ejemplos paradigmáticos[41]. Así, por ejemplo, implica una apertura a identificar criterios jurisprudenciales y a no olvidar que, con o sin EdP, la judicatura debe fundamentar sus decisiones lo que, por definición, excluye presentar la evaluación probatoria de la forma en que se venga la gana.

Antes de pasar a las conclusiones, veamos brevemente un caso acerca de cómo opera un EdP entendido como un prototipo. El caso ilustra, por una parte, que la metáfora del prototipo es satisfactoria para resolver el problema de la dispersión de las respuestas y, por la otra, cuestiona las ideas de Laudan acerca de las propiedades que aquellos debieran tener. A partir de la prueba disponible en el caso, resulta fácil construir una argumentación según la cual el MADR se entienda superado, pues vuelve innecesario seguir escarbando en otras posibles respuestas a la situación planteada. Para facilitar el análisis, recurriremos a un prototipo como el de las buenas novelas policíacas, pues aquel da cuenta de una clase de discurso explicativo que depende de la satisfacción del criterio de la plausibilidad de la reconstrucción elegida y de la muy escasa plausibilidad de reconstrucciones rivales[42]. Como todo prototipo, el de las novelas policíacas admite diferencias según sea el sujeto que lo aplique, pues la experiencia y evaluación de la literatura de ese género es variable. Aquello no obsta a que los desacuerdos serán limitados. Así, si alguien se excede en (o minusvalora) los requerimientos de completitud, de precisión o de coherencia, llegará un momento en que se le cuestionará su calidad de lector que entiende lo básico del género de las novelas policíacas y, por ende, si su prototipo es adecuado[43].

---

[41] Justificar un prototipo podría ser una operación similar a justificar una máxima de la experiencia. A juicio de Sandoval (2023, p. 259) justificar una máxima de la experiencia requiere desarrollar cuatro pasos: «(i) excluir conocimiento que no cumpla con las exigencias del derecho, (ii) incluir aquel conocimiento que sí responda a tales exigencias, (iii) precisar el conocimiento difuso y (iv) explicitar la máxima de forma que esta pueda llegar a evaluarse». Lo interesante es que, de forma similar, los prototipos deben justificarse de forma ad-hoc, exhibiendo las propiedades que posibilitan un juicio comparativo en cada caso. Nada excluye que estas propiedades puedan ser sometidas a escrutinio por parte de los intervinientes y sean efectivamente controladas cuando vulneren criterios de corrección propios del sistema de valoración probatoria.

[42] Por cierto, no queremos sostener que un buen prototipo del MADR sea coincidente con el de las buenas novelas policíacas. Lo usamos simplemente a modo de ilustración ya que, probablemente, todos los lectores de este texto disponen de uno (y bueno). Un buen prototipo del MADR debiese articular distintas propiedades esperables de las buenas sentencias judiciales como foco, precisión, coherencia, anclaje, economía e integridad discursiva, y claridad. Véase Coloma, (2025, pp. 195-239). La concurrencia de estas propiedades asegura que la reducción de la dispersión no se alcance por la vía de la adopción de decisiones sesgadas.

[43] Hay criterios, por cierto no uniformes, para distinguir un buen relato policíaco de uno malo. Chesterton, (2011, pp. 33-37, 87-91, 141-145).

El caso refiere al asesinato de Emilienne Gerard (E.G.) del cual fue condenado Louis Voisin (L.V.)[44]. El enigma surge cuando el tronco y las extremidades de la víctima son hallados dentro de cuatro sacos en un parque londinense. En el interior de uno de ellos hay un papel con la leyenda *Blodie Belgiam*, palabras que nada significan en inglés. La investigación comienza sin gran éxito hasta que un agente le pide a uno de los varios sospechosos (L.V.) que escriba la locución *Bloody Belgium* (Bélgica sangrienta) y este escribe *Blodie Belgiam*. Con posterioridad, se encuentra la cabeza de E.G. en un sótano al que se pudo acceder usando unas llaves en posesión de L.V. Junto a ello se produce un cúmulo de nuevas pruebas en contra de L.V. ¿Cuáles serían las claves para afirmar que en este caso se supera el MADR?

> *1)* La historia disponible, a partir de las pruebas, explica por qué L.V. habría querido dar muerte a E.G. Se demostró que: i) L.V. había terminado recientemente una relación sentimental con E.G., a quien le debía dinero; en poder de L.V. se encontraron dos libretas de ahorro de la víctima, etc.
>
> *2)* La historia disponible explica que L.V. pudo haber dado muerte a E.G. y que no se ha problematizado que dicha muerte y ocultamiento ocurrió (fue descuartizada). Así, se probó que: i) L.V. era carnicero y que era capaz de realizar los cortes que se le hicieron al cadáver; ii) que tenía las llaves del sótano donde encontraron la cabeza de E.G., etc.
>
> *3)* La historia disponible excluye a otras personas que potencialmente hubieren querido y podido dar muerte a E.G. según la manera en que se puede reconstruir lo ocurrido. Así, hubo evidencias de que: i) L.V. escribía *Blodie Belgiam* si se le pedía que escribiera *Bloody Belgium*; ii) un aro ensangrentado fue encontrado en una toalla en el departamento de L.V. y que hacía pareja con otro que fue encontrado en la oreja de la víctima, etc.[45].

¿En un caso así, el MADR requiere entenderse como umbral? No, pues no se está midiendo algo específico de las conexiones entre las pruebas y el hecho probado, sino que la relación se está comparando con la del prototipo[46]. Y no se piense que se trata de una metodología de trabajo inútil, pues al advertir las semejanzas con dicho prototipo se facilita la toma de una decisión. Así, no nos hemos quedado inmovilizados ante el riesgo de tomar una mala decisión, aun cuando reconocemos que en este caso aquel parece bajo[47]. A la vez, la aplicación del MADR contribuye a que no haya mucho ruido en el sistema (casos similares que tienen en cuenta el mismo prototipo debieran resolverse de forma similar).

¿Hemos entendido el MADR en clave objetiva? No necesariamente. La evaluación del nivel de conexión alcanzado entre la prueba rendida y la decisión probatoria adoptada, puede ser reconstruida en términos (parcialmente) subjetivos. Así, para aplicar el MADR se ha tenido en cuenta una comparación para la cual se ha tenido

---

[44] Véase Evans, (2006, pp. 65 y ss.); MacCormick, (1980. p. 49) y Coloma, (2025, pp. 121-126).

[45] En este caso además se produce una dinámica que es bastante propia de las novelas policíacas, cual es el chispazo de genialidad del agente de Scotland Yard al hacer escribir a L.V. la locución *Bloody Belgium*.

[46] La utilización del teorema de Bayes nos permitiría medir, pero ya hemos visto que Laudan desecha esa opción por la subjetividad que lleva implicada.

[47] Whitman, (2008, pp. 10 y ss).

que recurrir al prototipo de las buenas novelas policíacas (¿al declarar culpable a L.V. el lector de la sentencia queda en una situación, en alguna medida, semejante a la de quien lee una novela policíaca?[48]). Como hemos visto, el prototipo no es fijo sino depende de la experiencia específica del sujeto que lo aplica (no sabemos qué novelas policíacas ha leído o escuchado, ni cuáles considera buenas) y de la interacción con otros para desambiguar qué es una buena novela policíaca (hay intersubjetividad). Así, pese a toda la indeterminación que aquello pudiese implicar, no sería muy distinto a si, por ejemplo, se le pidiese que evalúe si la historia es plausible (es más, esto podría generar más dispersión). Un punto adicional que pudiese considerarse es que la estandarización podría ser mayor si a la comparación realizada se le añade un indicador subjetivo. Quien aplica el MADR podría considerar no sólo la semejanza con el género de novelas ya indicado, sino también, la persistencia en el tiempo del convencimiento de la culpabilidad de L.V., en los términos propios del prototipo utilizado (y que para formarse tuvo en consideración la conexión entre las pruebas y lo que da por probado)[49]. Esto refleja que en determinadas circunstancias lo cualitativo (que sería el juicio de semejanza respecto del prototipo), se puede complementar con lo cuantitativo (que sería la persistencia del punto de vista por un cierto período), y de esa manera funcionar mejor que si se les considerase aisladamente[50]. Así, la estabilidad de las creencias en una persona que se espera tenga una disposición a cambiar su punto de vista ante buenas razones (humildad epistémica) puede ser una buena guía para la decisión[51].

¿El EdP ha servido para distribuir errores? No, al menos, de manera directa o de manera exclusiva. En cuanto a esto último, contamos con razones para sostener que el MADR fue tenido en cuenta por Scotland Yard (o el Ministerio Público) para determinar cuánto tenía que investigar antes de llevar el caso a juicio, cuán minucioso debía ser en la fase de producción de prueba y de presentación de argumentos ante el tribunal, etc. Pero vamos a lo de la distribución de errores que es lo más interesante de la pregunta. El MADR no nos blinda del todo frente al riesgo de que el día de mañana aparezca nueva prueba que *eche por tierra* lo que, al día de hoy, es posible sostener respecto de la culpabilidad de L.V. (prestamos atención a esto porque es lo más cercano a los errores que preocupan a Laudan)[52]. Sin embargo, sí es más que suficiente para sostener que se trata de un discurso bien justificado. Lo que se diga en

---

[48] El uso de los EdP admite grados relativos al nivel de semejanza con el prototipo. Véase Clermont (2013, pp. 35 y ss.) y Coloma (2016a, pp. 52-54).

[49] Usamos esta reconstrucción del MADR no porque nos parezca especialmente acertada, sino simplemente porque para Laudan no constituye un EdP y, en el ejemplo, sí pareciera operar como tal.

[50] Laudan tiene razón en que la sola persistencia de la convicción es un mal estándar de prueba. Sin embargo, no es tanto por su subjetividad, sino por una falta de control de la creencia de base. En el ejemplo, esta última se controla por la aplicación del prototipo.

[51] Asumimos un nivel de sensatez básico para que funcione el EdP, pero Laudan también lo presupone, por ejemplo, para que evalúe la plausibilidad de una historia.

[52] Piénsese en el sorprendente y muy excepcional caso conocido como el regreso de Martin Guerre. Davis, (1984, pp. 1 y ss.).

la sentencia, nos situará en una posición que nos habilita para clausurar la discusión respecto de la culpabilidad de L.V., sin mucha dispersión respecto de casos similares[53]. Y más allá de que L.V. haya o no matado a E.G., el que se le haya condenado como tal resulta propio de un sistema de adjudicación que calificaríamos como más que decente.

Se nos podrá enrostrar que el caso elegido parece muy fácil, tal como lo anticipamos. Así, lo que esperamos del EdP no parece especialmente determinante como sí cabría esperar en muchos otros casos[54]. Lo importante es que, pese a ello, ilustra acerca de los usos de los EdP en que la concepción de Laudan pareciera no haber reparado. Así, a nuestro juicio, el ejemplo da cuenta de lo que son nuestras prácticas. Ellas, dentro de todo, alcanzan un nivel de operatividad, nada despreciable, para que la tarea judicial no se detenga ante los difíciles problemas propios de la toma de decisiones en contextos de incertidumbre.

5. CONCLUSIÓN

En este trabajo hemos analizado críticamente dos tesis y un presupuesto de Larry Laudan en torno a los EdP. Hemos formulado objeciones con el objetivo de enriquecer la discusión, explorar las limitaciones de su modelo y ofrecer una perspectiva que considere los límites y complejidades del razonamiento probatorio en la toma de decisiones judiciales.

En primer lugar, planteamos que, si bien las formulaciones subjetivas y ambiguas de los EdP podrían propiciar, *prima facie*, una mayor dispersión en las decisiones probatorias, aquello no es un obstáculo para sostener que se tratan, de todas formas, de EdP. Ello se debe a que no excluyen la estandarización ni el compromiso con conectar las pruebas con las hipótesis fácticas. Asimismo, como lo deja entrever la formulación de los EdP genuinos del propio Laudan, no parece plausible la formulación EdP objetivos *tout court* que sean consistentes con los propósitos que se espera que cumplan. Por cierto, las pretensiones de Laudan acerca de los procesos judiciales no están fuera de lugar si consideramos que existen exigencias que tienden a la reducción de la dispersión de las decisiones y a la construcción de conexiones entre pruebas rendidas y hechos probados. Nos referimos, como mínimo, a la deliberación en tribunales colegiados y a la obligación de motivar sentencias.

---

[53] Decimos esto, pensando que este caso hubiera de ser resuelto por jueces que hubieren tenido que justificar su decisión probatoria. En el caso del jurado, el foco debiera estar en la manera en que sería de esperar se hubiera desarrollado la deliberación.

[54] Sin perjuicio de lo señalado, creemos que una reconstrucción como la que hemos hecho sería útil para reducir malas decisiones probatorias en casos emblemáticos en que se debía aplicar el MADR. Nos referimos, por ejemplo, a que la reconstrucción que es posible hacer considerando a Sacco y Vanzetti (especialmente a este último) sería una muy mala novela policíaca. En cambio, la reconstrucción que es posible en el caso del asesinato de Nicole Brown y Ronald Goldman por parte de O.J. Simpson sí podría funcionar como una buena novela policíaca.

En segundo lugar, hemos examinado la relación entre los EdP y la distribución de falsos positivos y falsos negativos. La imposibilidad, en un alto número de casos, de reconocer que se ha incurrido en un error en la relación entre palabras y mundo, lleva a que la pretensión de que los EdP los distribuyan a través de decisiones judiciales la hace metodológicamente muy difícil de sostener. Sin embargo, este inconveniente no implica la inutilidad de los EdP. Ellos cumplen otras importantes funciones, como influir en la conducta de los distintos intervinientes en los procesos judiciales y validar los resultados obtenidos, lo que subraya el carácter constitutivo de las decisiones probatorias[55]. Finalmente, abordamos el problema de la dispersión excesiva en las decisiones en las cuales cabría esperar niveles limitados de desacuerdo, para lo cual hemos al concepto de «ruido».

En tercer lugar, a diferencia de los dos puntos anteriores, hemos abordado un presupuesto de Laudan consistente en considerar a los EdP exclusivamente como umbrales de suficiencia probatoria. Si bien atractiva para casos en que se podría llevar a cabo mediciones, la metáfora de los umbrales no permite dar cuenta de un universo amplio de casos que no requieren contabilizar la concurrencia de ciertas variables, sino realizar un análisis de semejanzas y diferencias. Recurrimos, entonces, a la metáfora de los prototipos para recoger formas argumentativas más consistentes con una evaluación desde la totalidad y que, en no pocas ocasiones, parece dialogar mejor con aquello de que la prueba en los procesos judiciales corresponde a una clase de asuntos de juicio en los que se trata de cautelar que los desacuerdos sean limitados.

Comenzamos este texto con un reconocimiento a la influencia ejercida por Larry Laudan en la discusión sobre los EdP. Su teoría (traducida al español) ha sido una de las principales responsables de una amplia literatura que de manera constante se está produciendo en relación con los EdP. La influencia de Laudan ha redundado en una búsqueda incesante de formas de dar operatividad a los EdP, la que, sin embargo, se ha visto aquejada por una dosis no menor de pesimismo. Vaya, entonces, nuestro decidido reconocimiento por la agudeza de Laudan al haber instalado con fuerza un problema, en el que muchos no habían reparado. El valor de aquello es más evidente al ir acompañado de un análisis lúcido y metodológicamente interesante. Nuestro análisis —que va de la mano de las ideas de otros autores— ha dejado, sin embargo, algunas *dudas razonables* acerca de dos de sus tesis y uno de sus presupuestos. Reconocer aquellas debilidades debiera ser un acicate para seguir explorando acerca de los usos de los EdP, pues —como creemos haber demostrado— *no todo está perdido*.

---

[55] Coloma y otros (2024, párr. 17) identifican como una "creencia-obstáculo" la resistencia a reconocer la dimensión constitutiva de la función jurisdiccional en la determinación de los hechos probados. Destacan que «no se trata de creencias necesariamente erróneas, sino de creencias que no resultan aptas para integrar de un modo coherente los elementos que son relevantes para describir una práctica».

BIBLIOGRAFÍA

Accatino, D. (2011). Certezas, dudas y propuestas en torno al estándar de la prueba penal. *Revista de derecho (Valparaíso)*, 37, pp. 483-511.

Agüero, C. (2016). Los estándares de prueba y el boom editorial del discurso probatorio en castellano. *Discusiones,* 18(2), pp. 81-106.

Aguilera, E. (2023). Los estándares de prueba: revisión crítica de su caracterización como umbrales o prototipos, y algunas alternativas. *Discusiones,* 30(1), pp. 227-274.

Azócar, M. J. (2018). ¿Cómo se construye la experticia judicial?: Una etnografía judicial en Tribunales de familia de Santiago de Chile. *Revista de derecho (Valdivia)*, 31(2), pp. 9-27.

Barrett, Lisa Feldman (2018). *La vida secreta del cerebro. Cómo se construyen las emociones.* Paidós (edición electrónica).

Bohm, D. (2018). *La totalidad y el orden implicado* (9a ed). Kairós.

Calderón, M. (2023). Acerca del irreductible ámbito de subjetividad en la formulación y aplicación de los estándares de prueba. *Quaestio Facti. Revista Internacional Sobre Razonamiento Probatorio*, 4, pp. 145-166.

Céspedes, T. (2021). Sobre la legitimidad del estándar de prueba en el proceso penal. *DOXA. Cuadernos De Filosofía Del Derecho*, 44, pp. 233-261.

Chesterton, G. (2011). *Cómo escribir relatos policíacos.* Acantilado.

Clermont, K. (2013). *Standards of Decision in Law*. Carolina Academic Press.

Coloma, R. (2016a). Los usos de los estándares de prueba: entre umbrales y prototipos. *Discusiones*, 18(2), pp. 23-57.

— (2016b). Algo más sobre umbrales y prototipos. *Discusiones*, 18(2), pp. 107-133.

— (2025). *Decidir hechos y argumentar con pruebas. Materiales para una teoría de la prueba en los procesos judiciales.* Marcial Pons.

Coloma, R., Agüero, C. y Lira, R. (2021) Tecnología para decidir hechos en procesos judiciales. *Revista Chilena de Derecho y Tecnología*, 10(1), pp. 111-143.

Coloma, R. Larroucau, J. y Páez, A (2024). Sobre el impacto judicial de la concepción racionalista de la prueba. *Revus [Online]*, *53*. Disponible en: https://journals.openedition.org/revus/10230.

Davis, N. (1984). El regreso de Martin Guerre. Antoni Bosch.

Dei Vecchi, D. (2020a). Prueba libre, justificación epistémica y el noble sueño de los estándares de prueba. *Revista de derecho (Valdivia),* 33(2), pp. 25-48.

— (2020b). Estándares de suficiencia probatoria, moralidad política y costos de error: el núcleo inconsistente de la epistemología jurídica de Larry Laudan. DOXA. Cuadernos De Filosofía Del Derecho, (43), pp. 397-426.

— (2022). Prueba sin convicción en su justa medida. *DOXA. Cuadernos de Filosofía del Derecho*, *45*, pp. 337-373.

— (2023). Sentencias judiciales con premisas fácticas falsas, *Revus [Online]*, 49. Disponible en: http://journals.openedition.org/revus/9406.

Evans. C. (2006). *The Father of Forensics* (capítulo III The Butcher of Soho). Versión Kindle.

Ferraris, M. (2023). *Documentalidad. Por qué es necesario dejar huellas.* Ediciones Trea.

Ferrer, J. (2007). *La valoración racional de la prueba*. Marcial Pons.

— (2025). *Prueba sin convicción: Estándares de prueba y debido proceso*. *Segunda edición actualizada y ampliada.* Marcial Pons.

Gama, R. (2016). A propósito de umbrales, prototipos y usos de los estándares de prueba. *Discusiones,* 18(2), pp. 59-79.

— (2020). Prueba y perspectiva de género. Un comentario crítico. *Quaestio Facti. Revista Internacional Sobre Razonamiento Probatorio*, *1*, pp. 285-298.

— (2022). En búsqueda de El Dorado. *Revus [Online]*, *43*. Disponible en: https://journals.openedition.org/revus/6773

Gascón, M. (2005). Sobre la posibilidad de formular estándares de prueba objetivos. *DOXA. Cuadernos De Filosofía Del Derecho*, *28*, pp. 127-139.
— (2010). *Los hechos en el derecho*. (3ra ed.). Marcial Pons.
Ginzburg, Carlo (2010). *El hilo y las huellas. Lo verdadero, lo falso, lo ficticio*. Fondo de Cultura Económica.
Ginzburg, Carlo (2013). *Mitos emblemas, indicios*. Gedisa.
González, D. (2020). ¿Es posible formular un estándar de prueba preciso y objetivo? Algunas dudas desde un enfoque argumentativo de la prueba. *Revista Telemática de Filosofía del Derecho*, *23*, pp. 79–97.
Guastini, R. (1999). *Distinguiendo*. Gedisa.
Gutting, G. (1980). The logic of invention. En T. Nickles (ed.) *Scientific Discovery, Logic and rationality*. (pp. 221-234). Dordrecht: D. Reidel.
Kahneman, D. Sibony, O. y Sunstein, C. (2021). *Ruido. Un fallo en el juicio humano*. Debate.
Kuhn, Th. (1993). La función de la medición en la física moderna. En T. Kuhn (ed.). *La tensión esencial*. (2ª reimpresión) (pp. 202 – 247) Fondo de Cultura Económica.
Laudan, L. (2005). Por qué un estándar de prueba subjetivo y ambiguo no es un estándar. *DOXA. Cuadernos De Filosofía Del Derecho*, *28*, pp. 95-113.
Laudan, L. (2013). *Verdad, error y proceso penal*. Marcial Pons.
MacCormick, N. (1980) The coherence of a case and the reasonableness of doubt. *The Liverpool Law Review*, spring-autumm, pp. 45-50.
Mazzarese, T. (1997). Scoperta vs. Giustificazione. Una distinzione molto illumninante o gravemente fuorviante. En L. Gianformaggio (ed.). *Scriti per Uberto Scarpelli*. (pp. 587-616) Giuffrè.
Páez, A. (2015). Estándares múltiples de prueba en medicina y derecho. En A. Páez (coord.), *Hechos, evidencia y estándares de prueba*. (pp. 123-152). Universidad de los Andes.
Sandoval, S. (2023). El sentido común y las máximas de la experiencia. *Ius et Praxis*, 29(1), pp. 250-256.
Schauer, F. (2004). *Las reglas en juego*. Marcial Pons.
— (2013). *Pensar como un abogado*. Marcial Pons.
Seth, A. (2023). *La creación del yo*. 3a ed. Sextopiso.
Taruffo, M. (2005). Tres observaciones sobre «Por qué un estándar de prueba subjetivo y ambiguo no es un estándar», de Larry Laudan. *DOXA. Cuadernos De Filosofía Del Derecho*, *28*, pp. 115-126.
Tuzet, G. (2021). *Filosofía de la prueba jurídica*. Marcial Pons.
Whitman, J. (2008). *The origins of reasonable doubt. Theological roots of the criminal trial*. Yale University Press.

Quaestio facti. Revista Internacional sobre Razonamiento Probatorio
Quaestio facti. International Journal on Evidential Legal Reasoning
Sección: Homenaje a Larry Laudan
2026 I 10 pp. 95-124
Madrid, 2025
DOI: 10.33115/udg_bib/qf.i10.23144
Marcial Pons Ediciones Jurídicas y Sociales
© Hylke Jellema
ISSN: 2604-6202
Recibido: 25/04/2025 | Aceptado: 0/11/2025 | Publicado *online*: 19/12/2025
Editado bajo licencia Reconocimiento 4.0 Internacional de Creative Commons

STANDARDS OF PROOF FOR FUTURE CRIMES AND DECISION THEORY

Hylke Jellema
Willem Pompe Institute for Criminal Law and Criminology,
Utrecht University

ABSTRACT: What is the proof standard for applying preventive criminal sanctions? This is an open question in various legal systems. Some authors suggest that we can answer it by using decision theory. On this approach, the proof standard is conceptualized as a probabilistic threshold: a preventive sanction can only be imposed if it is sufficiently probable that a person will commit a crime in the future. According to decision-theorists, how high this probability of a future crime should be, can be determined by means of a utilitarian calculus. However, such a decision-theoretic analysis requires wrestling with a number of difficult questions. This article surveys these questions and explores some avenues for answering them. It does so by considering a standard of proof for a fictional preventive sanction and offering a decision-theoretic justification for that standard.

KEYWORDS: Preventive sanctions; decision theory; legal proof; evidence theory; dangerousness.

SUMMARY: 1. INTRODUCTION.— 2. DECISION THEORY AND PROOF STANDARDS.— 3. A HYPOTHETICAL STANDARD OF PROOF.— 4. JUSTIFYING THE THRESHOLD: 4.1. Rejecting the empirical approach; 4.2. Reflective equilibrium; 4.3. Justifying the standard of proof; 4.4. Can we ignore correct outcomes?.— 5. THE VALUE OF AN IMPRECISE STANDARD: 5.1. The decision-theoretic argument for flexibility; 5.2. The need for a precise standard.— 6. PROOF STANDARDS AND RISK ASSESSMENTS: 6.1. Utilitarian assumptions in risk assessment technologies; 6.2. Proving a probability beyond a reasonable doubt.— 7. CONCLUSION.— REFERENCES

1. INTRODUCTION

In many countries, the (criminal) legal system is increasingly being used as a tool for preventing future crime (Carvalho, 2017). Part of this trend is the introduction of various criminal sanctions intended to prevent recidivism[1]. These are sanctions that can only be applied to a person if (among other things) they are deemed "dangerous" —i.e., likely to engage in criminal behavior in the future[2]. But when is someone sufficiently dangerous to be the subject of these sanctions? Another way of phrasing this question is "what is the proof standard for future crimes?" This is an open question in various legal systems (see e.g., Bijlsma & Meynen 2023; Tadros, 2013; Scurich, 2016; Schopp, 1996; Schopp & Quattrocchi, 1995; Slobogin, 1989; 2006).

One way of thinking about proof standards for dangerousness is as probabilistic thresholds that specify the minimum level of certainty that is required to prove a given factual statement. In the case of past crimes, such a factual statement may be "that the defendant committed the alleged criminal acts". For preventive sanctions, it could be "the defendant will (again) commit criminal acts in the near future". Or it may be a more specific proposition, such as "if not committed, the defendant will commit a violent act within the next year." But how high should we set the probabilistic threshold? One method for answering this question is *decision theory*. This is an approach to decision-making that is used in various fields, including economics, psychology and the law (Scurich & John, 2011, p. 90). It has also been extensively applied to legal standards of proof, especially the beyond a reasonable doubt standard (see e.g., Vorms & Hahn, 2021).

The underlying idea of the decision-theoretic approach to proof standards is that any proof standard should be set at the level where applying a (preventive) sanction maximizes expected utility. Where this point lies depends on the relative desirability (i.e., utilities) of the possible outcomes of a decision (not) to apply a (preventive) sanction. For instance, for past crimes the relative desirability of the possible outcomes is usually linked to Blackstone's (1962) remark that it is better that ten guilty persons escape than that one innocent suffer. In other words, a false conviction is at least ten times worse than a false acquittal[3]. This fact is commonly thought to jus-

[1] This is also known as risk-based sentencing (Eaglin, 2017).

[2] Whether someone is "dangerous" depends on the probability that they will cause future harm as well as on the magnitude of this harm (Scurich & John 2010, p. 446-7). Slobogin opts for a more fine-grained account, according to which "risk" comprises four elements: (i) a probability that (ii) a particular type of offence outcome will occur within (iii) a specific period of time (iv) in the absence of an intervention (Slobogin, 2021, p. 38).

[3] This ratio should not be taken too literally; various other ratios have been proposed. For examples, such as Voltaire's claim that "Tis much more Prudence to acquit two Persons, tho' actually guilty, than to pass Sentence of Condemnation on one that is virtuous and innocent" or Benjamin Franklin who wrote that "it is better a hundred guilty persons should escape than one innocent person should suffer" (Laudan, 2006, p. 63).

tify a high criminal standard of proof, somewhere between 90% and 99% (see e.g., Jellema, 2023, p. 102)[4]. After all, the higher we set the standard, the more difficult it will be to convict someone, and therefore fewer false convictions will occur. Yet a high standard also means that more people will be acquitted, including some who are guilty. Blackstone's remark implies that we are willing to tolerate a fair amount of such false acquittals to prevent a smaller number of false convictions and that the standard for proof beyond a reasonable doubt should therefore be high. For preventive sanctions, the assumptions that ought to guide a decision-theoretic analysis of the proof standard are less clear[5].

The literature on decision theory and proof standards has mostly been concerned with the beyond a reasonable doubt standard for past crimes. Nonetheless, there have been several authors who have applied the approach to preventive sanctions (Nagel, Neef & Schramm, 1977; Mossman, 1995; Scurich & John, 2010; 2011; 2012; Vars, 2012; Scurich, 2015; 2016; 2018). This body of work on decision theory and proof standards includes a number of scholars, such as Larry Laudan, who are skeptical of this approach (Laudan & Saunders, 2009). The result of this sprawling discussion has arguably been greater confusion rather than clarity. What we *can* conclude is that applying decision theory to proof standards means wrestling with a number of difficult philosophical, legal and practical questions. The goal of this article is to explore some of these questions as well as possible ways of answering them. In order to make this exploration as clear as possible, I use a fictional but realistic example of a preventive sanction and propose a proof standard for that sanction[6]. I then ask how one might justify this standard on decision-theoretical grounds, what questions this raises and how one might answer these questions.

The structure of this article is as follows. First, I explain the basics of decision theory (section 2). After that, I describe the preventive sanction and the associated standard of proof that will be used as a running example are introduced (section 3). Next, I offer a decision-theoretic justification of this proof standard (section 4). One of the key ideas developed in this section is that this justification can be reached by means of a "reflective equilibrium"—approach. Additionally, this part of the article also discusses the underexamined question whether we should include the utilities of correct outcomes when determining the proof standard, or whether we can limit ourselves to the utilities of errors. Section 5 deals with the question whether proof

[4] Though see section 3 for criticisms of the relationship between the Blackstonian ratio and the standard of proof.

[5] Though some suggest that a low proof standard is warranted because they consider a dangerous person who goes on to commit a crime (a false negative) worse than a preventive sanction being applied to a non-dangerous person (Monahan, 1977).

[6] The example used is fictional for two reasons. First, I do not mean to give the image that the aim of this article is to make a policy proposal about some existing preventive sanction. Rather, the goal is to explain decision theory. Second, the discussion in this article is philosophical in nature. I want to sidestep legal questions regarding criminal procedural law.

standards for future crimes can best be expressed in vague but flexible, qualitative terms or precise but inflexible probabilities. Finally, section 6 explores two questions regarding the relationship between predictions of dangerousness by experts and the fact-finder's (judge or jury) determination whether the proof standard is met. The first question is to what extent normative, utilitarian assumptions can be (or should be) made by the expert rather than the fact-finder. Second, I discuss the idea proposed by some authors and adopted in a number of legal systems that, aside from the probability of future crime, there should also be a proof standard for how confident we are in that probability.

2. DECISION THEORY AND PROOF STANDARDS

Decision theory is a method to balance costs and benefits in a decision dilemma. Before we turn to explaining this approach, it is helpful to distinguish it from a closely related idea with which it is sometimes confused. Many authors assume that a key aim of the beyond a reasonable doubt standard is to distribute errors fairly (see e.g., Jellema, 2023, p. 102). The underlying idea is that the higher we set the proof standard, the less likely it is that someone will be falsely convicted. However, this reduction in false convictions comes at a cost—the higher the proof standard, the more often defendants will be falsely acquitted. After all, the higher the proof standard is, the larger the group of defendants will be with strong evidence against them, but where the evidence is not sufficient to meet the standard of proof. The idea defended by various authors is that the proof standard should be set such that the criminal legal system produces roughly ten false acquittals for every false conviction (to the extent that such errors cannot be avoided). This ratio is based on Blackstone's (1962), that it is better that ten guilty persons escape than that one innocent suffer. This suggestion—that the proof standard ought to act as a method for distributing errors, has also been suggested in the context of preventive sanctions (see e.g., Monahan, 1977). The question for preventive sanctions would then be how many false positives (individuals subject to preventive sanctions, who would not have committed crimes in the future) are acceptable compared to every false negative (individuals to whom no preventive sanction was applied who went on to commit a crime).

While popular, this notion of error distribution has been the subject of critique. For example, as a number of authors note the distribution of errors depends on other factors aside from the proof standard (DeKay, 1996; Lillquist, 2002, Laudan & Saunders, 2009; Scurich & John, 2010). This includes unknowable factors such as how many defendants are actually innocent and guilty, the quality of the evidence against them and the degree to which courts or juries accurately evaluate this evidence. Furthermore, Laudan (2015) argues that (at least in the American criminal legal system) other criminal procedural rules, aside from the standard of proof influence the error distribution. In particular, there are numerous rules aimed at benefiting the defendant, for instance regarding the admissibility of evidence, thereby further shifting the

error distribution towards false acquittals. So, no matter how one sets the proof standard, we cannot guarantee that this will result in the desired error distribution.

In response to the aforementioned arguments, various scholars writing on the beyond a reasonable doubt standard suggest that Blackstone's ratio should not be taken as stipulating an optimal distribution of errors (see e.g., Lillquist, 2002; Laudan & Saunders, 2009). Instead, this ratio can also be interpreted as specifying the relative *utilities* of the two types of errors. This idea lies at the core of the decision-theoretic approach, according to which we can use decision theory to clarify standards of proof. The first to explore this approach were Kaplan (1967) and Cullison (1969), working independently. Their decision-theoretic analysis of the beyond a reasonable doubt standard begins with the assumption that the disutility of a false conviction is ten times greater than that of a false acquittal. These relative utilities can then be used to calculate how high this proof standard ought to be set so that the desirability of convicting would either be equal to or above the desirability of acquitting.

What would such a decision-theoretic analysis look like? The decision-theoretic framework relies on three types of variables: the courses of action one can decide to take, the probability that certain outcomes will materialize if a given course of action is chosen and the utilities associated with those outcomes. These utilities should not be thought of as the costs or benefits that an outcome *has* in some absolute sense. Rather, the notion of utility is simply a means of comparing how "good" and "bad" the potential outcomes are relative to one another (Scurich, 2016, p. 172; Scurich & John, 2010, p. 432).

To give a commonly used example: suppose that you have to decide whether to bring an umbrella when going outside or not. If you bring the umbrella, you would have to carry it, which is slightly cumbersome. In other words, bringing the umbrella involves a small amount of disutility. However, if you do not bring an umbrella and it rains, you will probably get wet, which you want to avoid. The disutility of getting wet is much greater than that of bringing the umbrella. So, should you bring it? Aside from the relative utilities of these outcomes, this also depends on the probability that these outcomes will come about. In this example, the most important probability is that of rain. If it is extremely unlikely that it will rain, it may not be worthwhile to bring an umbrella, because it is very likely that carrying it will not yield any benefits. If the chances of rain are very high, it will probably be in your best interest to bring it. It is then very likely that it will help you avoid getting wet. What you are trying to estimate is what choice—to bring or not bring an umbrella—maximizes your expected utility. The higher the probability of rain, the greater the expected utility gained by bringing the umbrella is. There is a certain point where the probability of rain is high enough that it is better to bring the umbrella than to not bring it. It is at that point that you should take your umbrella with you.

"If the probability of event X is at least probability P, take decision Y". The X and Y in this statement can be substituted for different kinds of events and decisions. This includes legal decisions. For instance, "if the probability that the defendant

committed the alleged criminal acts is at least 95%, convict them". This last sentence is one way to understand the proof of guilt beyond a reasonable doubt standard. As said, for criminal cases, most scholars agree that a false conviction is much worse than a false acquittal. And, according to many, this implies that the level of probability for proof of guilt beyond a reasonable doubt ought to be very high. For instance, according to Kaplan (1967), if we assume that a false conviction is ten times worse than a false acquittal, we end up with a proof standard of 91%. The formula Kaplan uses to arrive at this figure is as follows:

$$\text{Proof standard} = \frac{1}{1 + (U_{\text{false negative}} / U_{\text{false positive}})}$$

Where "U" standards for "utility". In the context of past crimes, a false negative is the acquittal of a guilty person. For future crime, it refers to a person to whom no preventive sanction was applied, who went on to commit a crime. In the context of past crimes, a false positive refers to an innocent person who is convicted. For future crime, a false positive occurs when we apply a preventive sanction to a person who would not have committed a future crime. When we take into account the relative utilities of these two types of errors, we arrive at a number between 0 and 1, expressing a proof standard between 0% and 100%.

While elegant, Kaplan's (1967) idea that proof standards can be determined by considering the utilities of errors has been criticized by various authors. The problem with his proposal is that there are more than two possible outcomes of a trial. More precisely, there are four:

Table 1 The four possible outcomes of the decision to convict or acquit

	Convicted	**Acquitted**
Guilty	True conviction (true positive)	False acquittal (false negative)
Innocent	False conviction (false positive)	True acquittal (false positive)

We can see that, aside from errors (false positives and negatives), a trial can also result in a true positive (the criminal conviction of a guilty person) or a true negative (the criminal acquittal of an innocent person). The same is true of preventive sanctions. Here the possible decisions are (i) applying a preventive sanction, or (ii) not applying such a sanction. Both decisions can be correct (true) or an error (false):

Table 2 The four possible outcomes of the decision (not) to apply a preventive sanction

	Sanction applied	**Sanction not applied**
To dangerous person	True positive	False negative
To non-dangerous person	False positive	True negative

For instance, a true positive occurs when we apply a preventive sanction to a person who would otherwise have reoffended (i.e., a "dangerous person"). A false negative means that no preventive sanction is applied, yet the person turned out to be dangerous and goes on to commit a crime. Yet while there are four possible outcomes, Kaplan's formula only takes into account two of them. As various authors have argued, decision theory is incomplete if it does not include the utilities associated with correct outcomes (e.g., Tribe, 1971; Lillquist, 2002; Laudan & Saunders, 2009; Nance, 2016, p. 24-5). To solve this problem, Tribe (1971) proposed a formula to calculate the standard of proof that includes all four outcomes:[7]

$$\text{Proof standard} = \frac{1}{1+\dfrac{(U_{\text{false negative}} / U_{\text{false positive}})}{(U_{\text{true negative}} - U_{\text{false positive}})}}$$

Including correct outcomes can have a significant effect on the resulting proof standard. For example, recall Blackstone's widely shared intuition that a false conviction is ten times worse than a false acquittal. Laudan & Saunders, (2009, p. 11-12) list various existing proposals that include this assumption but make different (reasonable) assumptions about the utilities of correct outcomes. Depending on which of these assumptions we adopt, the proof standard can be anywhere between 55% and 95%[8]. Let us consider one of these proposals to illustrate how the formula above works. Lillquist (2002, p. 109) offers the following example of how one could set the required utilities in his discussion of the beyond a reasonable doubt standard:

True acquittal: 0, True conviction: 1, False acquittal: -1, False conviction: -10.

He justifies these figures as follows: he assumes that there is no utility (positive or negative) for a true acquittal. He assigns a positive utility of 1 for true convictions, because it is the best possible outcome. Then he proposes that the downside of a false acquittal is that a guilty person is not convicted and that, hence, the negative social utility associated with this is the converse of the positive utility of a true conviction, i.e., -1. Finally, Lillquist assumes that a false conviction is ten times worse than a false acquittal, namely -10. Working from these assumptions, the resulting standard of proof would be 83%[9].

Because the utilities of correct outcomes can have a large impact on the resulting standard of proof there seems to be a good reason to include them when calculating this standard. However, there is also a downside to this. The equation that includes

---

[7] This formula was simplified by Laudan & Saunders (2009, p. 5) as follows:
Proof standard = $(U_{\text{true negative}} - U_{\text{false positive}}) / \{(U_{\text{true positive}} - U_{\text{false negative}}) + (U_{\text{true negative}} - U_{\text{false positive}})\}$

[8] Furthermore, they argue that certain plausible assumptions regarding these utilities could even lead to a standard of proof below 50% (Laudan & Saunders, 2009, p. 12).

[9] 1 / {1 + [(1 - -1) / (0 - -10)] } = 0.833. However, we can easily justify different numbers, that would result in a different standard of proof (Lillquist, 2002, p. 110).

all four outcomes makes it much harder to reach conclusions about the desired proof standard. For instance, if we want to conclude that the proof standard is high, it is no longer enough that the disutility of an erroneous conviction be much greater than the disutility of an erroneous acquittal. Instead the value of $U_{\text{true positive}} - U_{\text{false negative}}$ has to be greater than the value of $U_{\text{true negative}} - U_{\text{false positive}}$. Furthermore, as Laudan & Saunders (2009, p. 13-14) point out, the ratio I just mentioned is not a ratio of utilities, but a ratio of differences between utilities. One consequence of this is that we could, for example, subtract 10 from each of the utilities mentioned above by Lillquist (2002) and the result of the formula would be the same. However, in that case the ratios of the utilities would be quite different. For example, the utilities assigned to the two errors would be -11 for a false acquittal and -21 for a false conviction. These no longer stand in a 1 to 10 ratio to one another. In other words, utility ratios, such as Blackstone's ratio now tells us next to nothing at all about how high the proof standard ought to be. We can no longer just ask, "how much worse is a false positive compared to a false negative?" Instead, we have to answer the much more complicated question, what is greater, the utility difference between a true positive and a false negative, or the utility difference between a true negative and a false positive?

We now face a dilemma. On the one hand, it seems that we should not ignore correct outcomes, lest we end up with a proof standard that does not accurately reflect our utilitarian assumptions. On the other hand, including them means that it becomes much more difficult to determine how high the proof standard ought to be. Some authors attempt to circumvent this dilemma by making the simplifying assumption that the utilities of correct outcomes can be ignored. For instance, in the example of whether to take an umbrella, we did not consider the utilities of correct outcomes. This assumption is warranted if the utilities of both correct outcomes (taking the umbrella when it rains and not taking the umbrella when it does not rain) are roughly equal. In that case, Tribe's formula can be reduced to Kaplan's, where we only have to take into account the ratio of utilities of errors, a much simpler endeavor. As I argue in section 4, this solution is feasible for a decision-theoretic account of the proof standard for future crimes (while it may not be feasible for such an account of proof standards for past crimes).

There are further reasons why proof standards for preventive sanctions lend themselves more readily to decision-theoretic analysis than the standards for past crimes. First, decision theory involves a utilitarian calculus. Some may find this worrisome when it comes to decisions about past crimes, as these involve such unquantifiable considerations as justice, guilt and desert. Weighing these values against monetary costs, or against practical benefits such as deterrence may seem inappropriate. Yet preventive sanctions are part of the shift to preventive justice, where (criminal) law is used as a means to prevent future crime. This is an inherently utilitarian project (and is therefore also subject to the usual objections to the utilitarianism, such as that it ignores such values as justice). Second, decision theory involves analyzing proof standards in terms of quantitative probabilities. Tribe (1971) argued that such quantification leads to a number of practical difficulties, including whether jurors or courts

will be able to understand probabilistic instructions. However, dangerousness criteria involve predictions of the future, which is inherently a probabilistic assessment (whereas proof of past crimes is only probabilistic according to certain theories of evidential reasoning). This is all the more so because courts often base their judgments of danger based on structured risk assessment estimates experts that involve quantitative language (Vars, 2010, p. 873-4; Eaglin, 2017). Hence, this worry also seems less great when it comes to proof standards for future crime (Scurich & John, 2010).

Having explained the basics of decision theory, let us turn to the hypothetical preventive sanction and associated proof standard that will be used as a running example in this article.

3. A HYPOTHETICAL STANDARD OF PROOF

In this section I describe the proof standard that we will use as an illustrative example throughout this article. As said, this standard is intended to be fictional but realistic. Suppose that a criminal legal system allows for the following preventive sanction:

> *Treatment facility for repeat violent offenders:* Those who have been convicted of a violent crime on multiple occasions can be committed to a specialized institution to prevent recidivism. This commitment is for a period of up to two years. Within this detention center treatment options are offered, aimed at resocialisation.

Suppose that several legal criteria have to be met before this measure can be imposed. This includes a ´dangerousness´ criterion: there must be a danger that the felon will commit another violent offense if this preventive measure is not imposed. This dangerousness requirement is encapsulated in the following legal provision:

> *Proof standard:* The preventive sanction can only be applied if it is likely that the person will commit serious harm to others.

This proof standard specifies a minimum probability of violent recidivism. In this case, the word "likely" is intended to express a probability of around 50% that the person will commit certain acts. This required level of probability is justified using decision-theory in the next section. However, the legal provision does not contain a precise, quantitative probability, as the standard allows for some degree of flexibility in the minimum level of dangerousness required for applying a sanction. This flexibility is desirable on decision-theoretic grounds (see section 5).

4. JUSTIFYING THE THRESHOLD

The proof standard proposed in the previous section states that it must be "likely" that a person will commit serious harm to others. The word ´likely´ is a verbal expression of a probability. Such verbal expressions can, and often are, interpreted in

different ways within different contexts and by different people (see e.g., Willems, Albers & Smeets, 2020). In the current context the word "likely" is intended to convey a probability of recidivism around 50%[10], though the exact value may vary from case to case (see the next section). In this section, I explain how one could justify a standard of 50% for the fictional preventive sanction described in section 3. The overarching aim of this section is to show the kinds of questions one runs into during such a decision-theoretic justification and to explore some ways of answering them. In particular, one of the main ideas developed in this section is to propose a "reflective equilibrium"—approach to this process of justification (see 4.2).

4.1. Rejecting the empirical approach

The first step in our decision-theoretic analysis is to come up with utilities to put into Tribe's equation mentioned above. How do we do this? One common answer is that the proof standard for future crimes should be set according to the preferences of society in general or of a more specific subgroup such as lawmakers, psychiatrists or judges. What those preferences are is a question that can be, and has been, empirically investigated. Here are a few examples. Mossman & Hart (1993) investigated societal attitudes towards civil commitment. They asked participants which was worse, "being attacked by a man with a knife [a false negative, failing to commit a dangerous individual], or spending a certain time period as a patient in a state psychiatric hospital [a false positive, committing a non-dangerous individual]'. Similarly, Mossman (2006) asked mental health professionals to compare having to spend various lengths of time as a patient in state hospital to being attacked by a man wielding a knife. Scurich & John (2011) had two stakeholder groups, former mental patients and psychiatrists fill in a questionnaire about the use of restraints in psychiatric hospitals. They asked them which was worse and by how much, the unnecessary use of such restraints on patients (a false positive) or failing to use restraints when required (a false negative).

One problem with these studies is that they focus only on the costs associated with errors. As we saw in the previous section, this perspective may be too limited (however, see below for a defense of this approach). Laudan & Saunders (2009, p. 18) propose a more comprehensive approach to eliciting utilities comes in the context of the beyond a reasonable doubt standard for past crimes. According to Laudan and Saunders (2009), we can ask fact-finders (judges or jurors) how much they would pay to convert a false acquittal to a true conviction or a false conviction into a true acquittal. This approach would take all four outcomes into account.

[10] See Janus & Meehl who show that several US courts use the "likely" standard for civil commitments, but are typically unclear about what probability value they intend to convey with this. The authors tentatively propose that this standard can be translated into a minimum probability of recidivism of 50% (Janus & Meehl, 1997, p. 41).

Yet all these studies face a debilitating problem. If existing research shows anything, it is that people's intuitions about utilities differ wildly. For example, the studies of Mossman & Hart (1993) and of Mossman (2006) received such a variability in responses that summary statistics (e.g., the average or mean) would likely to have been misleading (Scurich, 2016, p. 175). The same problem appears in the study by Scurich & John (2011). They found that former mental health patients thought false positives were much worse than false negatives. The median patient tradeoff was that 141 false negatives were equivalent to one false positive. However, doctors thought that false negatives were worse than false positives. Their median tradeoff was seven false positives to one false negative. This suggests that personal experience may have an enormous impact on how one values the desirability of the outcomes.

These findings are in line with research on error ratios for past crimes, where great variability in responses was found (see Lillquist, 2002, p. 143-146). Furthermore, as Lillquist (2002, p. 144-5) argues, it may also be the case that people's responses would differ strongly if they had to make a decision in a real case, rather than being presented with an abstract, fictional case. However, if more detailed examples are used, it would be unclear how much the results would generalize. To summarize, there appears to be little hope for using empirical research to determine the societal consensus on the required utilities. Such a consensus appears simply not to exist.

4.2. Reflective equilibrium

The upshot of the above is that a decision-theoretic analysis of proof standards for future crimes cannot rely on a pre-existing social consensus regarding the relevant utilities. However, there is an alternative approach. In particular, we can propose a proof standard and show that the assumptions that underly this standard are reasonable.

One idea that can help us arrive at such a reasonable proof standard, resting on reasonable assumptions, is the philosophical *method of reflective equilibrium* (*cf.* Knight, 2023). On this approach, we begin by considering our opinions about some subject matter, then formulate systematic principles that account for these judgments. We then look for discrepancies between these principles and our judgments. If there are any, we adjust both principles and judgments until we reach a state in which they agree with one another[11]. I propose that we can use this approach to

[11] The version of the reflective equilibrium that I have in mind was already hinted at in the results found by Nagel, Lamm and Neef (1981, p. 368) who asked respondents about the desirability of false convictions compared to false acquittals and found that the reported values would sometimes lead to a surprisingly low standard for proof beyond a reasonable doubt. The authors go on to note that some respondents reevaluated their stated utilities after it was pointed out to them that their assigned utilities would result in such lenient proof standards. This reevaluation is in line with the notion of reflective

justify a proof standard for a preventive sanction. Specifically, my version of this approach for determining the standard of proof asks us to answer four questions:

1. What are the costs and benefits associated with the outcomes of the decision to (not) apply a preventive sanction?

2. What are the relative utilities of these outcomes given the results of step 1?

3. What is the resulting proof standard given the results of step 2?

4. Is this proof standard reasonable?

If the answer to the last question is "no"—i.e., if we have arrived at a standard that is unreasonably high or low, we return to steps 1 and 2, and ask whether we have overlooked any costs and benefits and whether we should change our assessment of the relevant utilities. If the answer to the final question is "yes", we end up with a clear justification for the assumptions that underly this standard. Those who wish to criticize this standard can then do so by questioning these assumptions.

4.3. Justifying the standard of proof

Let us apply the reflective equilibrium approach to the fictional preventive sanction and the associated proof standard described in section 3. The first step is to consider the associated costs and benefits associated with each outcome.

As Morris & Miller (1985, p. 23) argue, the two most important considerations when it comes to the utility of preventive sanctions are (i) how serious the interference with liberty involved is when the sanction is applied, and (ii) how serious the injury from the subsequent crime would be, were the sanction not applied to a dangerous person. To these two main considerations, several others can be added. For instance, being preventively detained because you are "dangerous" can have a stigmatizing effect, damage one's career, relationships and general health (Mossman, 1995, p. 111). Stevenson & Mayson (2022, p. 730) also mention the cost of detention as a cost of preventively detaining a person and as a potential cost of not applying such detention the harm to the victim's family and friends if a dangerous person does commit a crime. We can also draw inspiration from Laudan and Saunders (2009, p. 14-21) who list a number of utility-determining factors for true and false convictions and acquittals. Not all of these are applicable to preventive justice. For instance, preventive sanctions are not aimed at giving criminals their just deserts. Yet other factors mentioned by them are more relevant, such as "crime reduction by deterrence" and "crime reduction by incapacitation". Using these remarks as inspiration, I arrive at the following list of costs and benefits of each of the four outcomes (with arrows representing whether something is a cost or benefit and the number of arrows indicating of how great a cost or benefit it is):

---

equilibrium, as long as these respondents also updated their utility judgments to be coherent with the desired proof standard and if the resulting utility judgments remained reasonable to themselves.

Table 3: The costs and benefits associated with a decision (not) to apply the preventive sanction

	Correctly applied	**Incorrectly applied**
Sanction	*True positive*	*False positive*
	Deprivation of liberty ↓↓	Deprivation of liberty ↓↓
	Stigmatization (damage to career, personal relationships, health etc.) ↓↓	Stigmatization (damage to career, personal relationships, health etc.) ↓↓
	Cost of incarceration ↓	Person unjustly treated as dangerous ↓
	Crime reduction by deterrence ↑	Cost of incarceration ↓
	Crime reduction by incapacitation ↑	Crime reduction by dete-rrence ↑
	Sense of security community ↑	Sense of security community ↑
	Resocialisation ↑	
Non-sanction	*True negative*	*False negative*
	No costs or benefits, maintains status quo	Harm to victim(s) ↓↓↓
		Harm to family/friends victim(s) ↓
		Damage to sense of security community ↓↓

Many of these costs and benefits can be debated. For instance, is there indeed a distinct injustice in treating someone as dangerous when they are not? If so, should this be part of a utilitarian calculus, or is this strictly a deontological claim? My aim here is not to resolve such debates, but to explicate assumptions that will underly my decision-theoretic analysis in the hope that these will be reasonable to others (and to allow them to be criticized).

The second step is to consider the relative utilities of the four outcomes in the light of the above. A few assumptions seem reasonable. First, I assume that a true negative has a utility of 0, as it has no associated costs or benefits[12]. Second, for both true positives and false positives the associated costs are greater than the associated benefits. Hence, their utility is net negative[13]. Of these two, a false positive is worse (but not much worse) than a true positive. Finally, a false negative, where a dangerous person is not detained and hence commits a crime, is the worst outcome. More precisely, I follow Monahan's (1977, p. 370) suggestion that it is ten times worse than detaining a non-dangerous individual[14].

We can now arrive at numbers expressing the relative utilities of these numbers. Mossman (1995, p. 108-109) suggests setting the best outcome to 1 and the worst outcome to 0 in order to normalize the scale and to permit the easy evaluation of outcomes that fall between these extremes[15]. However, because three of the four outcomes have a net negative utility and because the best outcome has a net utility of 0, I will use a scale of 0 (best) to -1 (worst). Because a false negative is the worst outcome, it gets a value of -1. As just said, I assume that a false positive is ten times worse than a false negative, so the latter gets a value of -0.1. A true positive is better than a false positive. Hence, I assign it a value of -0.05. Finally, a true negative receives a value of 0.

If we enter these values into Tribe's formula mentioned in section 2. The result is as follows:

$$\text{Proof standard} = \frac{1}{1 + \{(-0.05 + 1) / (0 + 0.1)\}} \approx 0.1$$

In other words, the assumptions just made yield a proof standard of roughly 10%. The preventive sanction that we are using as an example can therefore only be applied to a person if there is a probability of at least 10% that they will commit a violent crime in the future. Yet this seems far too low. It would allow us to detain individuals even if there is only the slightest suggestion that they may soon commit

[12] Vars comes to the same conclusion. However, some might argue that not being labeled "dangerous" removes a stigma for a person and hence has some small amount of positive utility (Vars, 2012, p. 887).

[13] This may strike some as unreasonable for a true positive. Is it not a good thing to prevent a dangerous person from committing a crime? Yes, it certainly is. However, its goodness derives from the fact that it leads to avoiding a false negative and the associated costs mentioned in the table above. We should avoid double-counting such costs and benefits.

[14] One important consideration for preventive detention that was not included in this analysis is that, in practice, preventive detention is typically evaluated regularly. This dampens the potential harmfulness of false positives. After all, it creates the possibility of correcting such errors will be corrected within a relatively short time-period (Slobogin, 2018, p. 402).

[15] This also emphasizes that utilities are tools for comparing outcomes.

another crime. The resulting proof standard therefore fails the fourth step of the reflective equilibrium method.

We now have to return to steps one and two: should the costs and benefits and/or the utilities be changed? A candidate for revision immediately springs to mind. The assumption made above was that a false negative is ten times worse than a false positive (a reverse Blackstone ratio). This assumption was supported by citing Monahan (1977, p. 370). But let us look at what Monahan actually wrote:

> Paraphrasing Blackstone (1962), it might be better that ten "false positives" suffer commitment for three days than one "false negative" go free to kill someone during that period .

However, as Mossman (1995, p. 111-2) argues, "not all acts of violence, are murders. Some assaults frighten victims but cause little physical harm, and their evil approximates the evil of needless hospitalizations more closely than does an act of murder." Additionally, Monahan (1977) speaks of a hospitalization of three days, whereas the preventive sanction described in section 3 allows for detainment for up to two years. Hence, it seems reasonable to alter the ratio between false positives and negatives. Let us assume that a false negative is still worse in this case, but only slightly so. Because a false negative remains the worst option, it still has a value of -1. We can assign a value of -0.7 to a false positive to express this assumption that it is slightly less bad than the false negative. Finally, we shall retain the assumption that a true positive is half as bad as a false positive and that a true negative is a neutral outcome[16]. If we put these values into the equation, we get a proof standard of around 50%, which seems more reasonable[17]. In other words, we have arrived at a reflective equilibrium. While not all will find this standard reasonable, those who disagree with it can challenge the assumptions that underly it, namely that (i) a false negative is the worst possible outcome; (ii) a false positive is slightly less bad than a false negative, (iii) a true positive is less bad than a false positive, but still a negative outcome, (iv) a true negative is a neutral outcome.

4.4. Can we ignore correct outcomes?

As I explained in section 2, there are good reasons to include the utilities of correct outcomes in our calculation of the proof standard. Indeed, as we saw just now, true positives have various costs and benefits associated with them. Yet as I mentioned, the downside of including all four outcomes is that it makes calculating the proof standard significantly more complex. For this last reason, many authors make the simplifying assumption that the utilities of the correct outcomes are equal, and hence, cancel one another out (Lempert, 1976). This would mean that Tribe's formula can be reduced to Kaplan's, on which we simply compare the (dis)utilities of errors. Scurich (2016, p. 173-174) suggests that we can also make this assumption in

[16] Therefore, I assign these values of -0.35 and 0 respectively.

[17] $1 / \{1 + [(-0.35 + 1) / (0 + 0.7)] \} \approx 0.52$.

the context of preventive sanctions. However, he does not explore whether and why such an assumption might be reasonable. In fact, given the discussion above, one could argue that the two correct outcomes should not be regarded as equal. After all, we saw that true negatives have no associated benefits or costs, whereas true positives do have all kinds of costs and benefits. Nonetheless, I believe that an argument can be made for this simplification in the context of preventive sanctions.

To the extent that authors have tried to justify this assumption, they use one of two methods, which I call the "regret-based approach" and the "reductive approach". The regret-based approach is used by Lempert (1976, footnote 41). It is tied to the idea that the utilities one assigns in a decision-theoretic analysis can be viewed as the regret one may feel (or should feel) after making a decision. For example, a fact-finder might feel a certain amount of regret if they found out that they wrongfully applied a sanction to a non-dangerous person. This regret may be less than the regret they would feel if they wrongfully did not apply a sanction to a dangerous person. According to Lempert (1976), an ideal fact-finder should not feel regret after making a correct decision. Hence we can set the utility of both types of correct decisions to 0, thereby canceling them out.

Stevenson and Mayson (2022, p. 764) offer an argument against the regret-based approach. They note that ignoring the costs of correct decisions makes sense for adjudications of guilt, "where it is permissible to discount the harm inflicted on a person who is accurately convicted and punished because, at least in theory, that harm is deserved." However, this finding of guilt is absent in the preventive context. Therefore, Stevenson and Mayson (2022, p. 765) propose, even correct preventive decisions are costly because they "subordinate the welfare of the detained person to the public good."

The reductive approach is described (and rejected) by Laudan & Saunders (2009, p. 14-15). On this approach we do take the utilities of correct outcomes into account, but we do so by assigning negative versions of them to errors. For example, one benefit of a true conviction is that it gives a guilty person their just desert. We could say that a negative consequence of a false acquittal therefore is that it fails to give the party their just desert. If we strip away all costs and benefits of correct outcomes, we can set their utility to 0. However, as Laudan & Saunders (2009, p. 14-15) point out, this strategy does not work for proof of past crimes. They write:

> Consider the assignment of "a failure to deliver just deserts" to a false acquittal. But it is not only false acquittals that fail to give just deserts to the guilty. So do true acquittals and false convictions. We could readily add to the liabilities of a false acquittal that it too fails to give just deserts. Still, how do we capture [the fact that] a true acquittal fails to deliver an important benefit associated with a true conviction? The only reasonable way to do that is to reduce the utility of a true acquittal relative to a true conviction (or, alternatively, raise the utility of the latter). But either modification would undermine the whole enterprise, since the object of this maneuver was to render true convictions and true acquittals as neutered and thus dispensable, each possessing a utility of zero.

So, whereas the regret-based approach allows us to simply ignore all costs and benefits associated with correct outcomes, on the reduction account we must assign these costs and benefits to incorrect outcomes. In the case of convictions and acquittals, this account fails according to Laudan and Saunders (2009). But let us consider the aforementioned costs and benefits associated with preventive detainment. We saw that true negatives have no costs and benefits. Hence we can focus on the costs and benefits of true positives. Can we assign the converse of these to false negatives? I believe that we can. See the table below (converted costs and benefits in italics):

Table 4: the costs and benefits associated errors

	Incorrectly applied
Sanction	*False positive*
	Deprivation of liberty ↓ ↓
	Stigmatization ↓ ↓
	Person unjustly treated as dangerous ↓
	Cost of incarceration ↓
	Crime reduction by deterrence ↑
	Sense of security community ↑
Non-sanction	*False negative*
	Harm to victim(s) ↓ ↓ ↓
	Harm to family/friends victim(s) ↓
	Damage to sense of security community ↓
	No deprivation of liberty ↑ ↑
	No stigmatization ↑ ↑
	No cost of incarceration ↑↑
	No crime reduction by deterrence ↓
	No crime reduction by incapacitation ↓
	No resocialisation ↓

None of these conversions of costs or benefits appear to be problematic. So, it seems that we can reduce Tribe's formula to that of Kaplan:

$$\text{Proof standard} = \frac{1}{1 + (U_{\text{false negative}} / U_{\text{false positive}})}$$

To justify the 50% rule, we would have to conclude that a false negative is roughly equally bad as a false positive. This may seem surprising. Had we not said that a false negative is worse? However, note that we also assigned the converse of all attributes of a true positive to a false negative. Because I assumed that a true positive had a negative utility, we now add positive utility to the false negative. And because we

are simply moving the utilities from one part of the equation to the other, the result is the same.

One downside of this approach is that it makes determining the costs and benefits of the errors complex, because we have to convert each value to its negative. A benefit of the regret-based approach is that it is simpler in this regard. Rather than retaining the costs and benefits of correct outcomes in a different form, it ignores them [18]. As Stevenson and Mayson (2022) argue, however, this may count against this approach.

5. THE VALUE OF AN IMPRECISE STANDARD

The previous section showed how one can justify the proposed "likely" standard, which is intended to be a percentage around 50%. This justification was based on considering several costs and benefits associated with the different outcomes. However, these costs and benefits can differ between individual cases. For example, the type of violent crime that we expect to occur in the case of a false negative can differ in severity. Should the standard for the preventive sanctions we are considering therefore differ between cases too? This idea, that the proof standard should be flexible to account for the utilities in individual cases, has been argued for by several authors (Stoffelmayr & Diamond 2000; Lillquist, 2002; Vars, 2010). These authors claim that a quantified standard (i.e., one that states an explicit probability such as "50%") would not be flexible enough. Instead, they propose vaguer terms, of which "likely" is an example [19]. This section examines this suggestion in relation to the fictional preventive sanction central to this article. It will be argued that formulating standards of proof requires striking a balance between flexibility and normative guidance.

5.1. The decision-theoretic argument for flexibility

Vars (2010, p. 21) summarizes the point well, when he writes that the strongest argument against a quantified proof standard is that it is insufficiently flexible, as it "prohibits the trier of fact from balancing the costs of false positives and false negatives in a particular case and from adjusting the standard of proof accordingly." He therefore suggests using verbal rather than numerical standards. Similarly, Stoffelmayr & Diamond (2000, p. 783) write that a "single uniform standard across

[18] However, note that this means that the two approaches lead to different proof standards. For example, earlier, I wrote that I assigned a false positive a value of -0.7 and a false negative a value of -1. If we enter these values into Kaplan's formula, we get a proof standard of around 41%, rather than of 50%.

[19] Janus & Meehl (1997, p. 40) show that several US courts use the "likely" standard for civil commitments, but are typically unclear about what probability value they intend to convey with this.

cases is not an optimal resolution when the decisions to which the standard is being applied carry different costs." The most extensive treatment of the argument is given by Lillquist (2002) who proposes that the notorious vagueness of the beyond a reasonable doubt standard is desirable, as it allows fact-finders to use a flexible standard – tailored to the utilities of the individual case. Both Lillquist (2002, p. 88) and Stoffelmayr & Diamond (2000, p. 769) cite empirical research suggesting that, in practice, fact-finders do engage in such tailoring of the proof standard to the crime and to the punishment.

Lillquist (2002, p. 162-76) discusses the benefits and risks of having a flexible rather than a fixed standard for past crimes. The biggest problem with a flexible standard is that fact-finders may not accurately weigh the utilities in a given case and therefore apply a different standard than what is desirable. As Lillquist (2002, p.166 writes:

> It may be simpler for the juror to apply a rule that says "find the defendant guilty if you are ninety percent certain" than to apply a standard that, at best, implicitly invites the juror to decide what the appropriate standard should be[20].

He proposes two questions that we should ask when deciding whether variable or fixed standard of proof would be desirable: (i) is the preferred standard of proof likely to vary widely in different cases? (ii) are fact-finders likely to apply a standard of proof in a way that varies widely from the one that we would prefer in a given case? (Lillquist, 2002, p. 168). When we answer "yes" to the first question and "no" to the second, then a variable standard is preferable. If the answer to the first question is "no" and to the second "yes", we should prefer a fixed standard. He comes to the conclusion that a variable standard is preferable, in part on the ground that there can be very much variation in the utilities involved in individual criminal cases. Lillquist (2002, p. 170), points out that many legal systems have criminalized very minor crimes, such as manufacturing burglary tools, which are much less serious than, for example, murder and rape (and that there are many crimes in between these extremes, such as tax fraud). Similarly, the punishments can vary between a fine and life in prison (or even the death penalty in some legal systems) (Lillquist, 2002, p. 149-50). Therefore we would want a proof standard that can vary widely too.

What about proof standards for future crimes? As we saw earlier, people tend to have very different opinions on the utilities involved in preventive detention. On the other hand, while the beyond a reasonable doubt standard spans a wide range of cases and potential punishments, dangerousness criteria are often tied to a specific type of preventive sanction. The fictional sanction from section 3 is an example of

[20] Though applying a numerical threshold may not necessarily be easier for fact-finders. After all, judges and jurors have notorious difficulty with interpreting numerical information and find it difficult to critically question these numbers. This is also another reason why quantified standards may conflict with the fact-finder's task to consider the specifics of the case. As Tribe (1971, p. 1376) puts it, fact-finders may be "induced by the persuasive force of formulas and the precision of decimal points to perceive themselves as performing a largely mechanical and automatic role."

this, as it is connected to the treatment facility for repeat violent offenders. The range of cases covered by this standard should differ much less than those that fall under the beyond a reasonable doubt standard for past crimes. Nonetheless, as Mossman (1995, p. 110-1) writes:

> Not all acts of violence, however, are murders. Some assaults frighten victims but cause little physical harm, and their evil approximates the evil of needless hospitalizations more closely than does an act of murder. Moreover, involuntary hospitalization of persons who, if left alone, would have been harmless is more than a mere annoyance. It can be stigmatizing and damaging to an individual's career, marriage, and nervous system.

How much can the proof standard for our example differ if we adopt a variable standard of proof? I will consider two examples, that I take to be the most extreme cases. For the sake of simplicity, I will adopt the regret-based account discussed in the previous section and only consider the relative utilities of errors. The first example is based on Monahan's (1977, p. 370) suggestion that 'it might be better that ten "false positives" suffer commitment for three days than one "false negative" go free to kill someone during that period.' In other words, a relatively short period of commitment that will help prevent a murder. Let us therefore use Monahan's (1977) suggestion of a ratio of one false positive to ten false negatives as one of the extremes. An example from the opposite end of the spectrum would be inspired by Mossman's (1995) remark above, namely a long commitment of two years to prevent recidivism in the form of a violent assault that leaves the victim mostly unscathed. In the case of a false positive, a non-dangerous person is committed for two years. A false negative would mean that the aforementioned assault occurs. It does not seem unreasonable to reverse the ratio in this case, to one false positive for every ten false negatives.

When we put these values into Kaplan's formula, we get proof standards of 9% and 91% respectively. Of course one could argue that cases where a person who is likely to commit a murder only receives treatment for a brief moment are rare, as are cases in which someone who is not particularly dangerous will be committed for the maximum amount of time allowed. Nonetheless, even if we relax these assumptions, the point holds. For example, if we assume a 2:1 and 1:2 ratio the proof standard would vary between 33% and 66% respectively. So, it seems that we should answer Lillquist's first question with a resounding "yes".

As for his second question, Lillquist (2002, p. 174) admits that it is difficult to determine how often fact-finders would wrongly weigh the relative utilities. This is made all the more difficult because these utilities are dependent on the particulars of the case. Nonetheless, Lillquist (2002, p. 175) suggests that there is cause for optimism with respect to that question, in part because the reasonable doubt standard (with which he is concerned) has been vague for so long, and the legal system appears to be content with this vagueness. Yet, what about preventive sanctions? It is for this reason that I now turn to what I consider to be the strongest argument against a flexible standard.

5.2. The need for a precise standard

The argument against a vague, flexible standard is that it may leave fact-finders *too* free in their judgments. For example, as Stoffelmayr & Diamond (2000) argue, one criterium for an adequate standard of proof is that it should be precise enough to be distinguishable from other standards and that it should create at least some consistency (e.g., between jurors or judges in the same case). Yet as Kagehiro (1990, p. 194-197) found, jury instructions such as "very likely" and "extremely likely" all elicited essentially the same verdicts. In a similar vein, Morse (1982, p. 72) argues that standards for civil commitment are too vague and have no generally agreed upon meaning. He writes that each person therefore "injects his or her own private meaning into the criteria, rendering the system essentially lawless."[21] This point of lawlessness is also relevant in the criminal legal context. A well-known problem with preventive criminal justice compared to traditional criminal law, namely the lack of clear procedural rules and case-law that governs the process of proof (see e.g., Ashworth, Zedner & Tomlin, 2013; Bijlsma, 2024).

While variance in the standard of proof can sometimes be desirable from a decision-theoretic perspective, too much variance can lead to an unequal treatment of equal cases and a lack of predictability, which clashes with the requirement of legal certainty. Furthermore, as Tillers & Godfried (2006, p. 155) put it, triers of fact should not be permitted "to strike a balance that is wildly at variance with the values of society at large."[22] As we saw earlier, it is doubtful whether there truly exists a (measurable) societal consensus on the relevant utilities. Nonetheless, it is easy to imagine utility assignments that are deemed unreasonable by society (e.g., standards that are far too low or far too high). Finally, a precise standard would also make it easier to determine whether the fact-finder correctly applied the standard in individual cases.

According to Tillers & Godfried (2006, p. 155) it is up to lawmakers and appellate judges to select a standard of proof that best accommodates the competing interests[23]. They suggest that this should be a precise, quantitative standard. Yet there are also other options for the lawmaker or appellate court, which may strike a better balance between precision and flexibility. For instance, Stoffelmayr & Diamond (2000, p. 782) propose that "an alternative to [a] single probability standard (…)

---

[21] This claim is supported by the study of Monahan & Silver (2003) who asked judges how they would set proof standard for short-term involuntary civil commitment. They found values between 1% and 56%. See also (Slobogin, 2021, p. 48).

[22] Similarly, Scurich & John (2010, p. 446) argue that "any departure from the normative social cost policy should be justified by more than an appeal to judicial discretion" because otherwise it may be too easy to apply preventive sanctions, leading to an increase in the number of false positives.

[23] Similarly, Redmayne states that the weighing of utilities is a policy decision "better made by the legislature than by judges" (Redmayne, 1999, p. 183)

might be to provide a range (e.g., . 87 to .92) that [fact-finders] would be invited to apply according to their assessments of the costs of error associated with a particular offense." One objection to this proposal would be that, as we have seen, the proof standard for future crimes can vary wildly. An instruction that asks fact-finders to put the standard somewhere between "9% and 91%" (or even between 33% and 67%) is hardly helpful. Another approach might be to create more fine-grained standards, that limit the variance in cases (Stoffelmayr & Diamond, 2000, p. 782). For example, one could create multiple proof standards for the treatment facility for repeat violent offenders discussed in section 3, based on the utilities involved.

A further possibility suggested by Stoffelmayr & Diamond (2000, p. 783) is a non-quantitative standard that instructs the fact-finder to balance different costs and benefits. Lillquist (2002) argues that finding the right wording for this is extremely difficult. However, as Vars (2012, p. 893) points out, there are some courts that have formulated these types of jury instructions for dangerousness criteria. For example, one such instruction reads:

> In assessing the risk of reoffending, it is for the fact finder to determine what is "likely." Such a determination must be made on a case-by-case basis, by analyzing a number of factors, including the seriousness of the threatened harm, the relative certainty of the anticipated harm, and the possibility of successful intervention to prevent that harm.

This kind of instruction is aimed at jurors in common law systems. Yet in civil systems, where it is usually up to the judge to make such determinations, similar instructions could, for instance, be (and sometimes are) given by the lawmaker. Such an approach could also help avoid another potential problem with vague standards. We want to avoid situations in which the fact-finder includes considerations in their utility calculus that are socially undesirable (e.g., where they adopt a lower proof standard for certain social groups than for others). If the instructions were to include the factors that the fact-finder should weigh, this gives them guidance on how to go about determining the relative utilities in an acceptable way, without limiting their freedom to tailor the proof standard to the case at hand.

To summarize the above, the proof standard should be formulated in a way that balances the need for flexibility against the need for clear normative guidance to the fact-finder. The "likely" standard proposed here may allow too much flexibility. One could solve this by formulating several fine-grained quantitative standards, or by offering qualitative instructions on how the fact-finder should weigh the relevant utilities.

Does a proof standard that contains a precise probability requirement provide sufficient normative guidance? According to Slobogin (2021, ch. 2), it does not. He argues that policymakers should not only offer clarity about the probability of a future crime sufficient to justify preventive intervention, but also regarding the nature and time-frame of the predicted crime (Slobogin, 2021, p. 49-56). In addition, he opts for a subsidiarity requirement according to which a far-reaching intervention such as preventive detention is only permissible if there are no less intrusive measures

available to reduce the risk of a future crime to acceptable levels (Slobogin, 2021, p. 52-6). The proof standard used as an example in this article ("it must be likely that the person will commit serious harm to others") does not contain such details. Adding them would provide additional clarity for the fact-finder. However, note that the discussion above, regarding the balance between flexibility and normative guidance also holds for such additional criteria, where more normative guidance leads to less flexibility.

6. PROOF STANDARDS AND RISK ASSESSMENTS

On the decision-theoretic approach, the proof standard is conceptualized as a probabilistic threshold. In other words, to apply a preventive sanction, the probability of recidivism must be sufficiently high. The most important evidence for determining the probability of a future crime is often the testimony of an expert (typically a forensic psychiatrist) who uses a structured risk assessment instrument for determining the likelihood that the person under consideration will recidivate (Eaglin, 2017). Such instruments are increasingly "actuarial" (Vars, 2010, p. 873-4; Eaglin, 2017). This means that the expert uses statistical algorithms to combine and weigh various risk variables—such as age, gender and past violent behavior—and to combine these into a single probability judgment. The result of such an actuarial risk assessment may, for instance, be that a defendant is categorized as a "high risk individual", which might indicate a risk of violent recidivism of 60% within the next two years.

This section discusses two issues regarding the link between the expert's risk assessment and the fact-finder's decision whether the proof standard is met. The first is to what extent it can ever be up to the expert to weigh the utilities in a given case. The second is a suggestion made by some authors and adopted by some legal systems, that aside from a proof standard for the probability of a future crime, there should also be a proof standard for our confidence in this probability.

6.1. Utilitarian assumptions in risk assessment technologies

As we saw so far, on the decision-theoretic approach, we determine the level of the standard of proof by considering the utilities at stake. Furthermore, these utilities may differ between cases and hence so may the standard of proof. So far, I have assumed that the weighing of utilities is strictly up to the lawmaker or the appellate court and the fact-finder (i.e., the judge or jury, depending on the legal system). This fits with the literature on proof standards for past crimes. An expert is not supposed to offer verdicts on how to interpret the standard of proof or on whether a defendant's guilt is proven beyond a reasonable doubt. However, this is arguably different for dangerousness assessments. For those types of decisions, it is not unheard of for

the expert (or for the developers of the risk assessment technologies that the expert uses) to make normative, utilitarian assumptions. To give an example, Min & Ferris (2020, p. 14) discuss the Harm Assessment Risk Tool (HART). They write that this risk assessment technology is "calibrated to err on the side of caution, because it regards under-estimations of risk levels as a more serious error than over-estimations." In other words, the way in which the risk levels are reported is adjusted based on a normative belief about the required utilitarian tradeoff. As Eaglin (2017) argues, developers of risk-assessment tools make numerous of such normative judgments about when persons should be marked as "dangerous" or "high risk", including how much risk we should tolerate as a society and what tradeoff between false negatives and positives is desirable.

One reason why the fact-finder may not be aware of these normative choices made by experts and by the designers of risk assessment technologies is that risk assessments are often communicated to the fact-finder in categorical terms, such "low/medium/high risk", rather than in terms of probabilities. Scurich (2018) argues against this practice, suggesting that it obscures what is fundamentally a value judgment about the relative costs and benefits of correct and incorrect outcomes[24]. After all, whether a probability of recidivism of, say, 50% counts as high risk (which would, presumable justify applying a preventive sanction) depends on how one balances the utilities of the various outcomes. However, when risk assessments are communicated in categorical terms, the court may assume that the expert is using a different proof standard (based on different assumptions about the utilities involved) than they actually are (Bijlsma & Meynen, 2023, p. 273). This is especially problematic because courts typically follow the expert's judgments when it comes to whether an individual is "dangerous". When an individual is marked as "high risk", courts almost always opt for preventive interventions (Slobogin, 2021, p. 48).

A potential argument against leaving the balancing of utilities and the determination of the proof standard strictly up to the fact-finder is that a forensic expert may have what philosophers call "epistemic authority" (Zagzebski, 2012). In other words, the expert may be in a better position than the judge or jurors to assess what the consequences of applying the sanction to the defendant would be, as well as determining the consequences of recidivism. For example, they may have had in-depth conversations with the defendant when assessing how dangerous this person is and they may be able to draw on their own extensive experience with similar cases. Furthermore, it is likely that experts are better at interpreting the statistical evidence generated by actuarial risk assessment technologies. After all, many people find it difficult to interpret and draw correct inferences when reasoning about probabilities

[24] An even more neutral method of reporting risk assessments is offered by Slobogin. He proposes that experts should not report that a person has a specific probability of reoffending. Instead they should report only "that the offender received a risk score that is consistent with the scores of a group studied in previous research, X percent of which offended" (Slobogin, 2021, p. 45).

(Kahneman, 2011). Judges and jurors can also succumb to various kinds of errors in drawing conclusions for probabilistic evidence (Dahlman, 2024). Whether and to what extent the expert should make decision-theoretic determinations will depend on how we weigh their expertise on these matters against the lawmaker's democratic and the fact-finder's legal authority.

6.2. Proving a probability beyond a reasonable doubt

Some authors propose that, aside from a standard for the minimum probability of recidivism, there should also be a proof standard for the quality of our evidence for that probability. This idea was first discussed by Monahan & Wexler (1978). They make a distinction between two levels of proof standards. First there are those standards that measure the probability of future crimes, and that are expressed in terms such as "likely" or "highly likely" (I will call these "first-order standards of proof"). Second, there are standards that express how confident the fact-finder is about the evidence for this level of probability (which I will call "higher-order standards"). According to Monahan and Wexler (1978), such higher-order standards are expressed using terms like "clear and convincing evidence" and "beyond a reasonable doubt." They offer several real-life examples of proof standards for civil commitment requiring that a court find "beyond a reasonable doubt" that an individual is "likely" (or, in another example "more likely than not") to "do substantial harm to another" (Monahan & Wexler, 1978, p. 40)[25].

What would it mean to prove "beyond a reasonable doubt" that it is "likely" that the defendant will commit a violent crime in the future? There are at least two answers to this question. The first is offered by Monahan & Wexler (1978, p. 39) themselves. They note that risk assessment methods place an individual in a statistical group that has an associated risk of dangerous behavior. Suppose that membership of this group is based on having a certain age, sex, a psychiatric diagnosis, past criminal behavior and an addiction. We can verify all these facts about the person with a great degree of certainty, e.g., by checking their police records, bringing in witnesses and so on. In this example, proving a risk beyond a reasonable doubt means adducing solid evidence that the defendant falls into a given risk-category.

The second suggestion on what it means to prove a risk of recidivism beyond a reasonable doubt is explored by Vars (2012). He notes that estimates of risk themselves come with error – expressed in terms of confidence-intervals (Vars, 2012, p. 873-4; Van Koppen, 2008). Vars proposes that we can interpret higher-order proof standards in terms of being confident to a certain degree that the true probability of recidivism is *at least* a certain level. He gives the example of proving beyond a reasonable doubt – i.e., with a confidence of at least 90% – that there is a probabil-

[25] See Scurich & John (2010, p. 447-8) for further examples.

ity of at least 50% that an individual will recidivate (Vars, 2012, p. 876). According to Vars (2012), this 90% confidence can be achieved by requiring a higher predicted level of recidivism risk. The more the predicted level of recidivism risk exceeds 50%, the greater our confidence that the "true" recidivism risk lies above this point. How high the predicted level of risk ought to be to achieve this level of confidence depends on how accurate our risk-assessment methods are[26]. For instance, Vars (2012) discusses data concerning a real-life risk-assessment method. He concludes that for this method, 90% confidence that there is over 50% chance of recidivism for a given individual is achieved when the risk-assessment instrument predicts a risk of recidivism of 65% or higher (Vars, 2012, p. 876)[27].

Why would we want to have a two-tiered proof standard for preventive sanctions? One possibility is that such a standard reflects the idea that a preventive sanction such as the one used as an example here, can have far-reaching effects and that the decision to apply this sanction should not be taken lightly[28]. Even if our decision-theoretic analysis leads to a relatively low first-order standard, we may nonetheless want to be as confident as possible in our probabilistic assessment. This idea connects to a similar debate surrounding proof standards for past crimes: even if convicting a defendant maximizes expected utility, such a conviction may not always be warranted. In particular, many people have the intuition that we cannot convict someone based only on statistical evidence, even if this evidence makes it highly probable that they committed the alleged crime (*cf.* Dahlman, 2020; Günther, 2024a; 2024b). Some authors suggest that a conviction for a past crime is therefore only justified if there is a high probability of guilt and this probability is "stable"; (or "robust", "resilient" or "safe") (*cf.* Ho, 2008, 278; Stein, 2005, 88; Di Bello, 2013; 2015; Urbaniak, 2018; Günther 2024a; 2024b; Jellema, 2024). These authors have provided accounts of what it means for a probability to be "stable", why convictions based purely on statistical evidence are not stable and why convicting based on an unstable but high probability is undesirable. For future crimes, the statistical evidence (the output of a risk-assessment instrument) is typically the most important (and sometimes even the only) evidence based on which a preventive measure is applied. One unexplored question is whether the arguments from the debate on naked statistical evidence for past crimes also give reasons to desire "stable" probabilities for future crimes[29].

---

[26] I.e., it depends on how tight the associated confidence intervals are.

[27] In contrast Slobogin proposes that proving a recidivism risk of 51% to 90% certainty means that "the state would need to show that roughly 45 percent (51 percent x 90 percent) of people in the offender's risk category will commit 'a violent act constituting a threat to society'" (lobogin, 2021, p. 47).

[28] This idea might, for instance, underly the following remark by a court: "society has a substantial interest in the protection of its members from dangerous deviant sexual behavior. But when the stakes are so great for the individual facing commitment, proof of sexual dangerousness must be sufficient to produce the highest recognized degree of certitude" (Vars 2012, p. 868).

[29] A related question is whether (and when) proof of future crimes is based purely on statistical evidence. For this we'd need an account of the difference between statistical and individualized evidence. One such account is offered by Günther (2024a).

A potential, argument *against* a two-tiered standard of proof is offered by Slobogin (2006, p.144). As he puts it, the two-tiered standard of proof may be a "sleight of hand" because it "purports to require a high level of proof that the person will offend,' when in fact the conception really only considers the confidence in the prediction, rather than the risk". In other words, the suggestion of a two-tiered standard is nothing more than a method of masking a low standard of proof as high one. Another counterargument to such standards is offered by Laudan (2006, p. 119) regarding proof standards for past crimes. He proposes that the criminal legal system has a "pro-defendant bias". What he means by this is that various legal rules, such as those regarding the admissibility of evidence, are designed to benefit the defendant. According to Laudan (2006, p. 119-44) the intuition that underlies such pro-defendant rules may be mistaken. He argues that such rules increase the number of false acquittals above the desired level enshrined in the proof standard, thereby ignoring the negative utilities associated with that kind of error. A suggestion such as that by Vars (2012) may be subject to a similar critique. After all, on this account we assume that uncertainty about the probability that someone will commit a crime should benefit the defendant[30]. This could similarly be argued to be an example of the kind of "pro-defendant bias".

7. CONCLUSION

This article discussed how we may use decision theory to clarify standards of proof for future crimes. On this approach we weigh the utilities of the possible outcomes of the decision (not) to apply a preventive sanction. This utilitarian calculus yields a probabilistic threshold that has to be met before the preventive sanction can be applied. The principal aim of this article was to show that such a decision-theoretic analysis requires wrestling with a number of difficult questions and to explore some possible answers to them.

One of these questions was how we ought to determine the relevant utilities. This article rejected the suggestion that we can determine them by means of empirical studies. Instead, a "reflexive equilibrium approach" was proposed on which we use an iterative process to arrive at a reasonable standard of proof, based on clearly explicated assumptions. Another question that this article dealt with was whether our utilitarian calculus should include the utilities of correct outcomes. I tentatively suggested that this need not necessarily be the case and outlined two philosophical accounts that can support this conclusion (the "regret-based approach" and the "reductive approach").

[30] Another way of making this point is by noting that such so-called "higher-order uncertainty" about our "first-order probabilities" gives us just as much reason to revise our probabilities upwards as to revise them downwards (*cf.* Henderson, 2022).

A third question was how specific the proof standard for future crimes ought to be. Can it best be expressed in vague but flexible, qualitative terms or in terms of precise but inflexible probabilities? This article argued in favor of the former, but did note that the desired flexibility must be weighed against the normative guidance that this standard offers.

The final part of this article dealt with two questions regarding the relationship between predictions of dangerousness by experts and the fact-finder's determination whether the proof standard is met. First, to what extent can (or should) normative, utilitarian judgments be made by the expert (as often happens in practice)? Second, aside from a standard for the required probability of future crimes, should there also be a standard for the quality of the evidence on which that probability is based?

The aforementioned questions are part of the overarching issue where the limits of preventive justice lie. The aim of preventing future crime must always be weighed against the importance of individual liberty. Decision-theory is a method of formalizing and thereby clarifying this balancing act.

REFERENCES

Ashworth, A., Zedner, L., & Tomlin, P. (Eds.). (2013). *Prevention and the Limits of the Criminal Law*. Oxford University Press.

Bijlsma, J. (2024). Risicostrafrecht en rechtsorde: Kent noodzaak geen recht?. *University of Groningen.*

Bijlsma, J., & Meynen, G. (2023). Aannemelijkheid van strafuitsluitingsgronden en sanctievoorwaarden. Welke mate van onzekerheid is aanvaardbaar?. *Rechtsgeleerd Magazijn Themis*, *2023*(5), 268-276.

Blackstone, W. (1962). *Commentaries on the Laws of England* (Vol. 2). Beacon

Carvalho, H. (2017). *The preventive turn in criminal law*. Oxford University Press.

Cullison, A. D. (1969). Probability analysis of judicial fact-finding: A preliminary outline of the subjective approach. *Toledo Law Review,* 1: 538–598.

Dahlman, C. (2020). Naked statistical evidence and incentives for lawful conduct. *The International Journal of Evidence & Proof*, *24*(2), 162-179.

— (2024). A systematic account of probabilistic fallacies in legal fact-finding. *The International Journal of Evidence & Proof*, *28*(1), 45-64.

DeKay, M. L. (1996). The difference between Blackstone-like error ratios and probabilistic standards of proof. *Law & Social Inquiry*, 21(1), 95-132.

Di Bello, M. (2013). *Statistics and probability in criminal trials*. Stanford University.

Eaglin, J. M. (2017). Constructing recidivism risk. *Emory LJ*, *67*, 60-122.

Günther, M. (2024a). Legal proof should be justified belief of guilt. *Legal Theory*, *30*(3), 129-141.

Günther, M. (2024b). Probability of Guilt. *Canadian Journal of Philosophy*, *54*(3), 189-206.

Henderson, L. (2022). Higher-order evidence and losing one's conviction. *Noûs*, 56(3), 513-529.

Ho, H. L. (2008). *A philosophy of evidence law: Justice in the search for truth*. Oxford University Press on Demand.

Janus, E. S., & Meehl, P. E. (1997). Assessing the legal standard for predictions of dangerousness in sex offender commitment proceedings. *Psychology, Public Policy, and Law*, *3*(1), 33-64.

Jellema, H. (2023). (Im) probable stories: combining Bayesian and explanation-based accounts of rational criminal proof. *University of Groningen* [PhD thesis]

— (2024). Reasonable doubt, robust evidential probability and the unknown. *Criminal Law and Philosophy*, *18*(2), 451-470.
Kagehiro, D. K. (1990). Defining the standard of proof in jury instructions. *Psychological Science*, *1*(3), 194-200.
Kahneman, D. (2011). *Thinking, fast and slow*. Macmillan.
Kaplan, J. (1967). Decision theory and the factfinding process. *Stan L. Rev.*, *20*, 1065-1092.
Knight, C. (2023). Reflective Equilibrium, In: Zalta, E.N. & Nodelman, U. (eds.). The Stanford Encyclopedia of Philosophy. plato.stanford.edu/archives/spr2025/entries/reflective-equilibrium
Laudan & Saunders (2009). Re-thinking the criminal standard of proof: Seeking consensus about the utilities of trial outcomes. *International Commentary on Evidence*, 7(2), article 1 (online journal).
Laudan, L. (2006). *Truth, error, and criminal law: an essay in legal epistemology*. Cambridge University Press.
— (2015). Why Asymmetric Rules of Procedure Make It Impossible to Calculate a Rationally Warranted Standard of Proof. *Available at SSRN 2584658*.
Lempert, R. O. (1976). Modeling relevance. *Mich. L. Rev.*, *75*, 1021-1057.
Lillquist, E. (2002). Recasting reasonable doubt: Decision theory and the virtues of variability. *UC Davis L. Rev.*, *36*, 85-197.
Min, B., & Ferris, G. (2020). Regulating Artificial Intelligence for Use in Criminal Justice Systems in the EU. *Fair Trials,* 2-35.
Monahan, J. (1977). Strategies for an empirical analysis of the prediction of violence in emergency civil commitment. *Law and Human Behavior*, *1*(4), 363-371.
Monahan, J., & Silver, E. (2003). Judicial decision thresholds for violence risk management. *International Journal of Forensic Mental Health*, *2*(1), 1-6.
Monahan, J., & Wexler, D. B. (1978). A definite maybe: Proof and probability in civil commitment. *Law and Human Behavior*, *2*(1), 37-42.
Morris, N., & Miller, M. (1985). Predictions of dangerousness. *Crime and Justice*, *6*, 1-50.
Morse, S. J. (1982). A preference for liberty: The case against involuntary commitment of the mentally disordered. *Calif. L. Rev.*, *70*, 54-106.
Mossman, D. (1995). Dangerousness decisions: An essay on the mathematics of clinical violence prediction and involuntary hospitalization. *U. Chi. L. Sch. Roundtable*, *2*, 95-138.
— (2006). Critique of pure risk assessment or, Kant meets Tarasoff. *U. Cin. L. Rev.*, *75*, 523-610.
Mossman, D., & Hart, K. J. (1993). How bad is civil commitment? A study of attitudes toward violence and involuntary hospitalization. *The Bulletin of the American Academy of Psychiatry and the Law*, *21*(2), 181-194.
Nagel, S., Lamm, D., & Neef, M. (1981). Decision theory and juror decision-making. *The trial process*, 353-386.
Nagel, S., Neef, M., & Schramm, S. S. (1977). Decision Theory and the Pre-Trial Release Decision in Criminal Cases. *University of Miami Law Review*, *31*(5), 1433-1491.
Nance, D. A. (2016). *The burdens of proof*. Cambridge University Press.
Redmayne, M. (1999). Standards of proof in civil litigation. *Mod. L. Rev.*, *62*, 167-195.
Schopp, R. F., & Quattrocchi, M. R. (1995). Predicting the present: Expert testimony and civil commitment. *Behavioral Sciences & the Law*, *13*(2), 159-181.
Schopp, R. F. (1996). Communicating risk assessments: Accuracy, efficacy, and responsibility. *American Psychologist,* 51, 939-944.
Scurich, N. (2015). Criminal justice policy preferences: Blackstone ratios and the veil of ignorance. *Stanford Law & Policy Review*, *26*, 23-35.
— (2016). Structured risk assessment and legal decision-making. *Advances in Psychology and Law: Volume 1*, 159-183.
Scurich, N. (2018). The case against categorical risk estimates. *Behavioral Sciences & the Law, 36(5), 554-564.*
Scurich, N., & John, R. (2010). The normative threshold for psychiatric civil commitment. *Jurimetrics*, 425-452.

— (2011). Constraints on restraints: A signal detection analysis of the use of mechanical restraints on adult psychiatric inpatients. *S. Cal. Rev. L. & Soc. Just.*, *21*, 75-107.
Scurich, N., & John, R. S. (2012). Prescriptive approaches to communicating the risk of violence in actuarial risk assessment. *Psychology, Public Policy, and Law*, *18*(1), 50-78.
Slobogin, C. (1989). The ultimate issue issue. *Behav. Sci. & L.*, *7*, 259.
— (2006). *Minding justice: Laws that deprive people with mental disability of life and liberty*. Harvard University Press.
— (2018). Preventive justice: A paradigm in need of testing. *Behavioral sciences & the law*, *36*(4), 391-410.
— (2021). *Just algorithms: Using science to reduce incarceration and inform a jurisprudence of risk*. Cambridge University Press.
Stein, A. (2005). *Foundations of Evidence Law*. Oxford University Press.
Stevenson, M. T., & Mayson, S. G. (2022). Pretrial detention and the value of liberty. *Virginia Law Review*, *108*(3), 709-782.
Stoffelmayr, E., & Diamond, S. S. (2000). The conflict between precision and flexibility in explaining" beyond a reasonable doubt". *Psychology, Public Policy, and Law*, *6*(3), 769-787.
Tadros, V. (2013). Controlling risk. *Prevention and the Limits of the Criminal Law*, 133-55.
Tillers, P., & Gottfried, J. (2006). A collateral attack on the legal maxim that proof beyond a reasonable doubt is unquantifiable. *Law, Probability and Risk*, 5, 135–157.
Urbaniak, R. (2018). Narration in judiciary fact-finding: A probabilistic explication. *Artificial Intelligence and Law*, 1–32.
Van Koppen, P. J. (2008). De beschaving van risicostrafrecht: Tussen goede opsporing en prima gevaarspredictie. *Proces*, *2008*(2), 36-46.
Vars, F. E. (2010). Toward a general theory of standards of proof. *Cath. UL Rev.*, *60*, 1-45.
— (2012). Delineating sexual dangerousness. *Hous. L. Rev.*, *50*, 855-898.
Vorms, M., & Hahn, U. (2021). In the space of reasonable doubt. *Synthese*, *198* (Suppl 15), 3609-3633.
Willems, S., Albers, C., & Smeets, I. (2020). Variability in the interpretation of probability phrases used in Dutch news articles—a risk for miscommunication. *Journal of Science Communication*, *19*(2), A03.
Zagzebski, L. T. (2012). *Epistemic authority: A theory of trust, authority, and autonomy in belief*. Oxford University Press.

Quaestio facti. Revista Internacional sobre Razonamiento Probatorio
Quaestio facti. International Journal on Evidential Legal Reasoning
Sección: Homenaje a Larry Laudan
2026 I 10 pp. 125-156
Madrid, 2025
DOI: 10.33115/udg_bib/qf.i10.23147
Marcial Pons Ediciones Jurídicas y Sociales
© Edgar Aguilera*
ISSN: 2604-6202
Recibido: 29/04/2025 | Aceptado: 16/07/2025 | Publicado *online*: 04/11/2025
Editado bajo licencia Reconocimiento 4.0 Internacional de Creative Commons

¿FÁBRICA DE INOCENTES? ¡NO TAN DEPRISA! (DE UN POSIBLE SESGO EN EL PROYECTO DE EPISTEMOLOGÍA JURÍDICA DE LAUDAN)

Edgar Aguilera*
Becario posdoctoral en el Barcelona Institute of Analytic Philosophy (BIAP), Unidad de Excelencia María de Maeztu; sede: Universidad de Girona
edgar.aguilera@udg.edu

RESUMEN: Luego de aportar el contexto en el que surge la metáfora de la fábrica de inocentes (sección 1), y de vincularla con los diagnósticos general y específico que Laudan hace respecto del sistema acusatorio estadounidense y similares (secciones 2 y 3), se ofrecen elementos que sustentan la tesis de que, en la base del proyecto de epistemología jurídica del autor, yace un sesgo confirmatorio que lo lleva a sobredimensionar la frecuencia y costos de los falsos negativos (secciones 4 a 7). A lo anterior le sigue la identificación del posible origen del sesgo referido (sección 8). Se cierra con la exposición sucinta a) de una teoría alterna de la justificación del castigo basada en la comunicación moral que se entabla con la persona condenada, y b) de un proyecto de epistemología jurídica compatible con esta (sección 9).

PALABRAS CLAVE: sistemas acusatorios; epistemología jurídica; veredictos erróneos; sesgo de confirmación; teorías de la justificación del castigo.

---

* Agradezco la minuciosa revisión y sugerencias de las personas dictaminadoras anónimas y, muy especialmente, a la profesora Amalia Amaya y al profesor Andrés Páez por su amable invitación a participar en el I Mini-foro de Epistemología Aplicada: Homenaje a Larry Laudan, llevado a cabo en agosto de 2023 en el Instituto de Investigaciones Filosóficas de la UNAM, México. En mi intervención en dicho evento se encuentra ya el gérmen de las ideas aquí presentadas. Un agradecimiento también muy especial al profesor Víctor García Yzaguirre por su invitación a que presentara una versión previa de esta investigación en la sesión inaugural del seminario interinstitucional «Perspectivas sobre la teoría del derecho», a los asistentes a ese seminario y, por supuesto, al profesor Raymundo Gama, encargado de preparar un comentario formal para aquel encuentro.

MANUFACTURED INNOCENCE? NOT SO FAST! (A POSSIBLE BIAS IN LAUDAN´S LEGAL EPISTEMOLOGY)

ABSTRACT: After providing the reader with the context of the metaphor appearing in the article´s title (section 1), and linking it with the general and specific diagnoses that Laudan offers regarding the American accusatory system and others relevantly similar (sections 2 and 3), I put forward some elements in support of the thesis that a confirmation bias underlies the author´s legal epistemology project; a bias leading him to overstate the frequency and costs of false negatives (sections 4 to 7). I go on to identify the possible origin of the abovementioned bias (section 8), and to conclude I sketch a) an alternative theory of what justifies the infliction of legal punishment (which is centered around the notion that a moral communication is established between the State and those convicted of a crime), and b) a legal epistemology compatible with such a theory (section 9).

KEYWORDS: accusatory systems; legal epistemology; erroneous verdicts; confirmation bias; theories of legal punishment.

SUMARIO: 1. EL CONTEXTO DE LA METÁFORA DE LA FÁBRICA DE INOCENTES.— 2. EL RIESGO QUE, SEGÚN LAUDAN, CORREN LOS SISTEMAS ACUSATORIOS Y SU DIAGNÓSTICO GENERAL RESPECTO DEL SISTEMA ESTADOUNIDENSE Y SIMILARES: APLICANDO LA METÁFORA.— 3. UN DIAGNÓSTICO MÁS ESPECÍFICO (¿Y CERTERO?) DE PARTE DE LAUDAN.— 4. ALERTA SOBRE UN POSIBLE SESGO.— 5. RÉQUIEM POR EL NÚCLEO DURO (O HARDCORE) DE LA EPISTEMOLOGÍA JURÍDICA.— 6. LA DILUCIÓN DE LA ESTRUCTURA ACUSATORIA TRIPARTITA.— 7. EN NOMBRE DE LA CIENCIA, LA FILOSOFÍA Y LA EVIDENCIA EMPÍRICA.— 8. UNA PECULIAR TEORÍA UTILITARISTA DE LA JUSTIFICACIÓN DEL CASTIGO.— 9. REFLEXIÓN DE CIERRE: ESBOZO DE UNA TEORÍA ALTERNATIVA DE LA JUSTIFICACIÓN DEL CASTIGO Y DE UN PROGRAMA DE EPISTEMOLOGÍA JURÍDICA BASADO EN ELLA. — BIBLIOGRAFÍA.

1. EL CONTEXTO DE LA METÁFORA DE LA FÁBRICA DE INOCENTES

A finales de la primera década de este siglo, los periodistas franceses Anne Vigna y Alain Devalpo (2010) publican en castellano el libro *Fábrica de culpables: Florence Cassez y otros casos de la (in)justicia mexicana*. En sus páginas, los autores buscan sumarse a la denuncia de lo que los más optimistas dirían que *era* uno de los problemas estructurales del sistema mexicano de administración de justicia penal, a saber: el débil sustento probatorio que, no obstante, con frecuencia alcanza para condenar[1]. Tan

[1] O alcanzaba, si se comparte el optimismo referido.

débil sería ese sustento[2] que, en lugar de la excepción, inculpar falsamente[3] podría ser la regla, según lo sugiere la obra referida. De ahí el título de «fábrica de culpables».

En virtud de lo arraigado de su «modelo autoritario de investigación» (Magaloni, 2009), de su justicia «solo de escritorio» (Zepeda, 2004), y de la incompetencia, corrupción y abusos a los que, por décadas, se acostumbraron sus «investigadores de papel» (Azaola y Ruíz, 2009), no comparto el optimismo de quienes sostendrían que en México lo anterior ya no es un problema, —si es que, de hecho, alguna vez lo fue— o, al menos, no uno del que haya que preocuparse demasiado[4]. Sin embargo, entiendo que hay razones que pueden conducir a que se piense así. La principal en mi opinión tiene que ver con la célebre reforma de 2008, que hizo transitar a la administración de justicia mexicana de una pauta preponderantemente *inquisitoria* a otra con énfasis *acusatorio*[5].

La pregunta, claro, es ¿por qué este cambio podría dar pie a un optimismo de esa naturaleza? La respuesta en buena medida se relaciona con el aura positiva que envuelve a los sistemas acusatorios paradigmáticos, con el buen nombre del que en principio gozan, mismo que no está del todo injustificado. Y es que, en efecto, estos sistemas suelen acoger en su estructura un conjunto de dispositivos procesales explícitamente orientados a condicionar la emisión de una condena a la superación por parte de las fiscalías del duro reto de tener que identificar, aportar y practicar *pruebas de la mayor contundencia* en contra de los acusados. La idea fuerza[6] es que, como base del ejercicio de la potestad punitiva del Estado haya una investigación más robusta, un esfuerzo demostrativo de la culpabilidad mucho más considerable que el que cabría anticipar en ausencia de los dispositivos aludidos.

La esperanza es que, al complicar de ese modo la procedencia de una condena, (al menos de la que resulta de un juicio[7]) simultáneamente se reduzcan de forma

---

[2] Aunado a las no pocas violaciones al debido proceso en las que las autoridades suelen incurrir al recabar los elementos que conforman los acervos probatorios respectivos.

[3] Es decir, condenar al inocente.

[4] Para reducir la dosis de especulación con la que se le aborda, la cuestión de determinar la frecuencia de las condenas falsas (o *wrongful convictions*) tendría que zanjarse empíricamente, en la medida de lo posible. Eso es precisamente lo que se ha intentado hacer mediante iniciativas como la de los denominados registros de exoneraciones. El que ha llevado la batuta es el The National Registry of Exonerations, véase: *https://exonerationregistry.org/*; al que se han sumado el European Registry of Exonerations, véase: *https://www.registryofexonerations.eu/*; el Miscarriages of Justice Registry de Reino Unido, véase: *https://evidencebasedjustice.exeter.ac.uk/miscarriages-of-justice-registry/*; entre otros.

[5] Al menos en el papel, y siguiendo con ello la tendencia marcada por los Estados Unidos a los países de la región.

[6] Una de ellas, al menos.

[7] Hago esta aclaración, pues es común que nos olvidemos de las condenas resultantes de la conocida como «justicia negociada», cuya dinámica —si no incluye un estándar de prueba mínimamente riguroso, y/o si se le otorga al fiscal una amplia discreción para modificar los cargos imputados con el propósito de que el acusado acepte al menos alguno de ellos, y/o para proponer la imposición de penas mínimas en contraste con las que podrían ser impuestas por el juzgador, etcétera—, puede subvertir las buenas intenciones de dificultar la procedencia de condenas en la fase del juicio. Ese olvido es más

drástica las probabilidades de que se castigue injustamente. ¿Qué bases hay para que se tenga esa esperanza? Que, después de todo, no parece que sea frecuente que, en contra de alguien realmente inocente, puedan recabarse pruebas de tal envergadura[8], excluyendo, por supuesto, que se trate de un montaje deliberado que involucre la siembra o manipulación de aquellas.

Por su parte, entre los dispositivos de los que se espera que dificulten la obtención de una condena (y que reduzcan, por tanto, el riesgo de condenar falsamente) destacan la presunción de inocencia, la carga impuesta al Estado de probar la culpabilidad, y un umbral de suficiencia probatoria muy elevado, que típicamente es invocado mediante la expresión de «prueba más allá de toda duda razonable» (de que el acusado cometió el delito respectivo).

2. EL RIESGO QUE, SEGÚN LAUDAN, CORREN LOS SISTEMAS ACUSATORIOS Y SU DIAGNÓSTICO GENERAL RESPECTO DEL SISTEMA ESTADOUNIDENSE Y SIMILARES: APLICANDO LA METÁFORA

Habiendo establecido sus aparentes bondades, el problema con los sistemas de corte acusatorio[9] es que, de acuerdo con Laudan, pueden verse capturados por una suerte de *garantismo exacerbado* o de *hipergarantismo* que, en su afán de evitar a toda costa las condenas falsas, está dispuesto:

1) A conceder a las personas acusadas cuanta ventaja procesal y/o probatoria que sea posible concebir, —yendo así más allá (y, en sus versiones extremas, mucho más allá) de la inclusión de los dispositivos procesales antes enlistados— y, por tanto;

2) A aceptar con indiferencia la consecuencia previsible de una concesión de ese tipo: el crecimiento exponencial de las absoluciones jurídicamente válidas[10], pero (probablemente) falsas[11].

En otros términos, si en el trabajo de Vigna y Devalpo (2010) antes aludido se habla de una *fábrica de culpables*, de la mano de una actitud hipergarantista y a la manera de un efecto de péndulo, un sistema acusatorio, podría afirmar Laudan, co-

---

que curioso si se considera que a la justicia negociada parece recurrirse de modo preponderante en las distintas regiones en que operan los sistemas acusatorios.

[8] Suponiendo que el alto nivel de contundencia probatoria al que se aspira haya sido formulado de forma adecuada en términos de no verse afectado de un elevado grado de indeterminación y también en términos de referirse fundamentalmente a la capacidad justificativa del conjunto de pruebas.

[9] Cuya instauración, o la de algo aproximado a ellos, —al menos en lo que se refiere a la estructura acusatoria tripartita (conformada por la presunción de inocencia, la carga del Estado de probar la culpabilidad y la exigencia de un estándar probatorio sumamente riguroso)—, es considerada ya un compromiso ineludible del Estado democrático y liberal.

[10] Si suponemos que están basadas en la correcta determinación de que el estándar de prueba respectivo no fue satisfecho.

[11] En el sentido de liberar acusados materialmente culpables.

rre el riesgo de situarse bruscamente en el extremo opuesto: en el de operar también como una *fábrica,* pero *de inocentes*. De inocentes en su mayoría tan artificiales como los culpables que produce su metafórica contraparte fabril, pero sumamente dañinos para la sociedad si, como sugiere el autor, se toman en cuenta los delitos futuros que, según él, es razonable predecir que serán cometidos —por la proporción no menor de ellos con antecedentes penales—, durante el periodo en prisión al que podían haber sido condenados, de no ser por el efecto de las múltiples e inapropiadas concesiones que esta fábrica de inocentes hace a favor de todo acusado.

Para Laudan, lo anterior no es solo un riesgo abstracto, sino que es precisamente lo que sostiene que ha ocurrido ni más ni menos que con el sistema acusatorio estadounidense, es decir, el haberse transformado paulatinamente en lo que, por usar uno de los modos de caracterizarlo al que el propio autor recurre, puede ser descrito como un sistema «excesivamente amigable con la absolución» y, por extensión, sumamente tolerante con la comisión de falsos negativos[12].

Por supuesto, dicho diagnóstico alcanzaría a cualquier otro sistema acusatorio considerablemente semejante, es decir, que incluya la tradicional estructura acusatoria tripartita conformada por la presunción de inocencia, la carga del Estado de probar la culpabilidad, y un exigente estándar de prueba; pero que, además, entre otras cosas[13], insista en la implementación de un estándar probatorio aún más severo, incorpore reglas que directamente ordenen la exclusión de pruebas de cargo relevantes, o que de forma indirecta tengan ese efecto, y/o que impida la apelación de los veredictos absolutorios[14].

---

[12] Esta caracterización del sistema de justicia estadounidense puede apreciarse mejor en la versión en inglés del primer libro en que Laudan reflexiona sobre el proceso penal desde una perspectiva epistemológica. Refiriéndose a las ventajas de posibilitar que las absoluciones sean apeladas, el autor sostiene que «[…] *apart from enabling the state to convict more of the truly guilty than it now does*, opening up the system to appeals of acquittals would halt the drift of evidence law in the direction of becoming ever more acquittal-friendly. *If appellate courts were routinely confronting injustices of both sorts, we have every reason to expect that the courts would come up with fixes that worked to reduce the total number of erroneous verdicts* instead of producing remedies that focus myopically on false convictions […]» (énfasis añadido, Laudan, 2006: 211).

[13] Y en esas otras cosas adicionales radica, según Laudan, lo exacerbado del garantismo.

[14] Para apreciar más vívidamente la preocupación que en Laudan genera lo que aquí he llamado el hipergarantismo, reproduzco el siguiente largo pero muy representativo pasaje: «Consideremos por un momento la situación de un acusado inocente que enfrenta un juicio oral. Él sabe… que, *gracias al (elevado) estándar de prueba aplicable, corre solo un pequeño riesgo de ser condenado incorrectamente*… Sabe, además, que aun si es condenado *tiene la oportunidad de que su condena se revoque en apelación*. También está enterado de que, gracias a la presunción de inocencia, *el jurado debe ignorar el hecho de que muchos de los anteriores actores en su drama (policías, fiscales y jueces) recabaron pruebas significativas en su contra o consideraron que las hay*. Sabe asimismo que, al menos que opte por invocar una defensa afirmativa, *no está obligado a decir o a hacer nada para lograr su absolución*… Hasta aquí, sus intereses coinciden con los de la sociedad. No obstante, pese al riesgo relativamente menor antes mencionado, puede perfectamente entenderse que nuestro hipotético inocente, desafortunadamente acusado de un delito, siga preocupado por la pequeña pero innegable posibilidad de ser condenado y, además, de que su condena errónea sea confirmada en la apelación respectiva. Así las cosas, a él (y mayormente a los

3. UN DIAGNÓSTICO MÁS ESPECÍFICO (¿Y CERTERO?) DE PARTE DE LAUDAN

En su más reciente y desafortunadamente último libro sobre cuestiones epistemológico-jurídicas, Laudan (2016) se da a la tarea de ofrecer un diagnóstico más específico respecto del sistema acusatorio estadounidense. Y lo hace centrándose en los denominados «crímenes violentos»[15] que, de acuerdo con la estadística pertinente, fueron cometidos en aquel país durante el 2008.

Pues bien, según Laudan (2016: 5-6), en ese año se cometieron alrededor de 1.7 millones de dichos crímenes. En respuesta, la policía arrestó a 595 mil sospechosos[16]. Por su parte, ya sea por iniciativa del fiscal o de un juez, se desestimaron los cargos contra 217 mil arrestados. De los restantes, 333 mil fueron condenados por vía de un acuerdo (o «plea bargain agreement»), dejando con ello alrededor de solo 45 mil personas cuyo caso prosiguió a la etapa de juicio oral, de las cuales 30 mil fueron condenadas y 15 mil fueron absueltas. Consideradas en conjunto hubo entonces 363 mil personas condenadas, ya sea mediante acuerdo o luego del juicio respectivo, y las restantes 232 mil personas arrestadas salieron libres (ya sea por la desestimación de los cargos en su contra o por haber recibido una absolución).

El siguiente paso en su proceder fue incorporar la crucial información relativa a la frecuencia y costos asociados tanto a los falsos positivos[17] (Laudan, 2016: 50-56,

---

abogados defensores en general) *le gustaría minimizar la probabilidad de ser condenado injustamente aún más*. Pues bien, partamos de que en efecto existen ciertas pruebas de cargo considerables, pues de otra manera su caso no hubiese llegado tan lejos. Siendo así, nada le complacería más que *descubrir que hay un conjunto de reglas de exclusión que podrían tener el efecto de evitar que el jurado tenga contacto con algunas de aquellas pruebas*. En este punto, *su interés en ser absuelto y el interés de los genuinamente culpables en que a ellos también se les absuelva, comienzan a converger*; al tiempo que el interés de aquellos comprometidos con la búsqueda de la verdad, como nosotros, empieza a discrepar del ferviente deseo de librarse de una condena que es plausible decir que tienen ambas clases de acusados (los genuinamente inocentes, por supuesto, *pero también los genuinamente culpables*). ¿Por qué esa discrepancia? Porque, aunque indudablemente nos gustaría ver absuelto al acusado inocente, *comprendemos perfectamente que toda concesión probatoria otorgada a este también beneficiará a los materialmente culpables, y a veces, mucho más a estos últimos*. Por tanto, *una vez que mediante el estándar de prueba se han reducido los riesgos de una condena falsa para el acusado inocente, debemos oponernos a que generosamente se sigan repartiendo más y más ventajas a favor de los acusados, pues sabemos que muchos de los que se beneficiarán de esa condescendencia serán los genuinamente culpables, y que, si ponemos más y más obstáculos para la admisión de pruebas relevantes, será cada vez más difícil que con las que sobrevivan se les pueda condenar*» (énfasis añadido, Laudan, 2013: 183).

[15] Esta categoría comprende al homicidio doloso o *murder*, la violación o *rape*, el asalto agravado o *aggravated assault* y el robo a mano armada o *armed robbery*.

[16] A 22 mil por homicidio doloso, a 36 mil por violación, a 288 mil por asalto agravado y a 158 mil por robo a mano armada.

[17] Expresión que en el vocabulario de Laudan comprende a los erróneamente condenados, ya sea como resultado de un juicio o de un acuerdo.

73-78), como a los falsos negativos[18] (pp. 56-66, 68-73). En este orden de ideas, según los cálculos de Laudan, el número estimado de falsos positivos en el periodo contemplado fue de 10.9 mil[19]. Su costo, por otra parte, ha de entenderse en términos de las víctimas que dichos falsos positivos producen, cantidad que asciende a 24 mil. ¿Por qué esta cifra?

Porque Laudan piensa que la cantidad de falsos positivos debe multiplicarse por 2.2. Este último factor es el resultado de sumar dos cifras que representan a las víctimas producidas por esa clase de veredicto erróneo. Por un lado, tenemos al inocente condenado por cada falso positivo, al que Laudan considera en sí mismo una víctima[20], o sea, la cifra de 1; y, por otro, tenemos la cantidad de víctimas de los crímenes violentos que, con base en los datos estadísticos consultados por el autor, serán cometidos en promedio al año por el verdadero aunque desconocido perpetrador, en cuyo lugar fue castigado el inocente que cada condena falsa implica (asumiendo, claro, que dicho perpetrador es un ofensor reincidente). Ese promedio es de 1.2 crímenes violentos anuales, cifra de la cual Laudan extrae la misma cantidad de víctimas (o sea, 1.2)[21].

En lo que toca a los falsos negativos, es decir, a los culpables liberados, su número estimado en el periodo estudiado asciende a 93.8 mil. Laudan obtiene esta cifra de considerar que el 38 por 100 de las 217 mil desestimaciones de cargos son erróneas[22], aunado al 80 por 100 de errores en las 15 mil absoluciones registradas. Por su parte, el costo en víctimas asociado a la totalidad de falsos negativos es de 112 mil. Dicha

---

[18] Esta expresión comprende a los culpables materiales liberados como resultado de la desestimación de los cargos en su contra o por haber recibido una absolución (al término de la fase de juicio oral).

[19] Cifra a la que se llega aplicándose una tasa del 3 por 100 de condenas falsas, que es la tasa en la que, en términos generales, según el autor, coinciden los estudios empíricos más robustos en materia de exoneraciones logradas por los múltiples proyectos de inocencia activos en los Estados Unidos.

[20] Entonces, para Laudan, tenemos, de un lado, a la víctima del delito; delito que quedó impune (aunque en su momento la víctima no lo sabe y quizá nunca lo sepa), pues no se identificó al verdadero perpetrador, sino que un inocente «tomó su lugar». Y, de otro, a la víctima del falso positivo, es decir, al inocente condenado. Para Laudan, este es tan víctima como quien fue objeto del delito respectivo. ¿Por qué? Porque si suponemos que el castigo que se impone a los considerados culpables de un delito es semejante o proporcional, en términos del sufrimiento infligido, al daño ocasionado por el acto delictivo mismo, entonces ambas víctimas experimentan, digamos, un dolor o pesar equivalente. ¿Por qué entonces no incluir a la víctima del acto delictivo en la cuantificación del costo de una condena falsa si a fin de cuentas estamos hablando de víctimas? Porque Laudan se centra exclusivamente en las víctimas que puede decirse que son causadas, directamente en el caso de los falsos positivos e indirectamente en el caso de los falsos negativos, por los veredictos erróneos, más no por los actos delictivos mismos. El caso de los falsos negativos es especial porque, en efecto, en el cálculo de sus costos Laudan considera fundamentalmente a las víctimas de los actos delictivos que muy probablemente llevarán a cabo los ofensores que ya son reincidentes o que son, por usar una expresión del autor, criminales de carrera. Sin embargo, siguiendo al autor, esos actos y sus correspondientes víctimas, no habrían tenido lugar si los perpetradores con perfil reincidente no hubiesen sido absueltos por un sistema que le brinda a los acusados ventajas excesivas. De modo que, en algún sentido (indirecto he dicho antes), son esas absoluciones falsas las causantes del «desastre».

[21] De modo que, sumando 1 y 1.2, se obtiene el factor de 2.2 por el que tiene que multiplicarse la cantidad de falsos positivos a los efectos de estimar sus costos.

[22] Erróneas en el sentido de liberar culpables materiales.

cifra resulta de multiplicar el número de culpables liberados, 93.8 mil, por 1.2, que corresponde al número de víctimas que cada falso negativo produce al dejar libre al materialmente culpable y reincidente que, como antes se dijo, en promedio al año cometerá esa misma cantidad, 1.2, de crímenes violentos.

A la luz de la información anterior, nótese primero, nos pide Laudan, la respuesta a todas luces insuficiente de parte del sistema al procesar solo 595 mil del más de millón y medio de crímenes violentos cometidos durante el año en cuestión. Por si eso no fuera ya escandaloso, nótense las 232 mil personas que se libraron de una condena, de las cuales, de acuerdo con el autor, a 93.8 mil puede considerárseles erróneamente puestas en libertad (en el sentido de tratarse de personas materialmente culpables). Nótese también que el costo de esos falsos negativos es de 112 mil víctimas y compáresele con las solo 24 mil víctimas atribuibles a los 10.9 mil falsos positivos.

Por otro lado, si dirigimos la mirada a la forma en que los errores se distribuyen, tenemos que esto ocurre a razón de alrededor de 8.6 falsos negativos (virtualmente 9) por cada falso positivo. Se trata entonces de una ratio bastante cercana a la supuestamente sugerida por Blackstone cuando sostuvo que es preferible que 10 culpables sean absueltos a que un inocente sea condenado. Por tanto, se trata de una ratio íntimamente ligada al estándar probatorio de «más allá de toda duda razonable».

¿Por qué el vínculo entre la ratio de Blackstone de 10 a 1 y el estándar mencionado? Porque la interpretación probabilística que a este se le suele dar —según la cual, para poder darse por probada, la hipótesis de culpabilidad debe haber alcanzado el 90 por 100 (o el 0,9) de probabilidades de ser verdadera (con base en las pruebas aportadas)—, es la que supuestamente tendrá el efecto de (o al menos contribuirá a) que los errores se distribuyan de ese modo, o sea, a razón de 10 absoluciones falsas por cada condena falsa.

Ahora bien, ¿por qué para Laudan es importante llamar nuestra atención a la casi equivalencia entre la ratio obtenida en su análisis (8.6/1) y la ratio de Blackstone (10/1)? Porque al haber mostrado el daño *concreto* que, según él, esa distribución del error causa —recuérdese, 112 mil víctimas atribuibles a las 93.8 mil absoluciones falsas contra las solo 24 mil víctimas atribuibles a las 10.9 mil condenas falsas—, Laudan piensa haber mostrado simultáneamente lo irracional que es mantener vigente el estándar de prueba que produce o de menos contribuye importantemente a la obtención de la distribución referida entre falsos negativos y falsos positivos, a saber, el estándar de «prueba más allá de toda duda razonable». La crítica es, en síntesis, que la dosis de tolerancia hacia las absoluciones falsas inserta en esa ratio (y en ese estándar) es simplemente inaceptable.

Si esto es así, parece entonces no quedar opción más que rebajar dicho estándar, sustituirlo por uno menos exigente, por ejemplo, por el estándar de «prueba clara y convincente»[23]. Por supuesto, es previsible que esta medida, en efecto, genere algunos

[23] Que es el estándar probatorio que, en el sistema estadounidense, supuestamente se sitúa a medio camino entre el de «la preponderancia de las pruebas» y el de «más allá de toda duda razonable».

falsos positivos más y, por ende, más víctimas atribuibles a dicha clase de error. No obstante, sostendría Laudan, eso se vería más que compensado con la sensible reducción tanto de falsos negativos, como de las víctimas atribuibles a estos. Y es precisamente eso, morigerar el estándar, cambiarlo por uno menos severo, al menos para el caso de los crímenes violentos (o delitos graves), lo que propone (Laudan, 2016: 88-109).

En adición, el autor sugiere otras medidas como limitar (y, si se pudiera, eliminar) las reglas que directamente ordenan (o indirectamente tienen el efecto de) excluir ciertas pruebas de cargo relevantes, sentar las bases para que las absoluciones puedan apelarse, requerir que los jueces se abstengan de solicitar al jurado que no valore negativamente la falta de cooperación del acusado durante la fase investigativa y/o su silencio durante el juicio, etcétera (pp. 110-137).

4. ALERTA SOBRE UN POSIBLE SESGO

Múltiples y muy variadas han sido las críticas dirigidas 1) a la metodología seguida por Laudan para la elaboración del que llamé *diagnóstico específico* en la sección previa, y 2) a sus propuestas para lidiar con la metástasis de absoluciones falsas de la que supuestamente padece el sistema estadounidense[24]. Y su reclamo de fondo es la precipitación con la que el autor parece haber procedido; señalamiento que comparto.

De modo más específico, pienso que Laudan incurrió constantemente[25] en una de las modalidades de error inferencial que él mismo advierte que se pueden cometer en el intento de esclarecer alguna cuestión empírica de interés, y que *grosso modo* consiste en darles a las pruebas o datos de los que se dispone un valor distinto del que racionalmente les correspondería, ya sea mayor o menor (Laudan, 2013: 38). En este sentido, Laudan parece haber sobrevalorado la información que corroboraba su intuición previa (a cuyo origen volveré después) de que los falsos negativos debían ser muy frecuentes y costosos, mientras subvaloraba la que pudiera contradecirla[26]. Por

---

En clave probabilística, el estándar de la preponderancia equivaldría a requerir que, para ser declarada probada, una hipótesis tiene que ser probablemente verdadera en más del 50 por 100 (0,5+); el de la prueba clara y convincente requeriría que las probabilidades de ser verdadera de una hipótesis sean del 70 por 100 (0,7); y el de la duda razonable, que las probabilidades de ser verdadera de la hipótesis de culpabilidad del acusado sean del 90 por 100 (0.9).

[24] Entre ellas, véanse las incluídas en el volumen 48, número 4, de la *Seton Hall Law Review*: Koppi, 2018; Findley, 2018; Zalman, 2018; Cassell, 2018. También las de Gardiner, 2017, y Dei Vecchi, 2020a, p. 145-182; 2020b.

[25] Y desde muy temprano en su reflexión epistemológico-jurídica, como intentaré mostrar más adelante.

[26] En este orden de ideas, una de las principales críticas de Findley (2018) es que Laudan se apresura al considerar correctas, sin más, las opiniones de los jueces, recabadas en el más importante estudio del siglo xx sobre el jurado estadounidense: el de Kalven y Zeisel (1966); estudio que fue clave en la estimación que Laudan hace de la supuestamente muy elevada frecuencia con que se pronuncian absoluciones falsas en su libro de 2016. El estudio de Kalven y Zeisel reporta que en los casos de la muestra respectiva en los cuales el jurado terminó absolviendo, pero los jueces habrían condenado, los jueces

opinan que, no obstante, había bases probatorias para condenar; o, dicho de otro modo, que el estándar PMATDR había sido satisfecho, y que, por tanto, se trataba en su gran mayoría de absoluciones, además de inválidas, muy probablemente falsas. A esta gran conclusión se le conoce como la «hipótesis de la liberación», a la cual volveré después con más detalle. Pues bien, Findley hace notar que el que no se tengan al menos algunas dudas con respecto a estas opiniones de los jueces ignora la gran lección de las condenas erróneas que a la fecha se sabe que han ocurrido. ¿Cuál? Que, en virtualmente todos los casos de personas exoneradas (gracias a los esfuerzos de algún proyecto de inocencia), las pruebas recabadas en su contra fueron suficientes para convencer a policías, fiscales, jurado, jueces y, a veces, a sus propios abogados defensores de que eran culpables conforme al estándar PMATDR (por eso fueron condenados en primer lugar). Teniéndose esto en cuenta, debería haber al menos algún grado de escepticismo saludable hacia las opiniones recabadas por Kalven y Zeisel, no solo de parte de Laudan, sino de la sociedad en general. Laudan parece ir en la dirección de ese escepticismo saludable, pues en efecto toma en cuenta uno de los más destacados estudios modernos acerca del jurado estadounidense: el elaborado por Givelber y Farrell (2012), en el cual, contrario al estudio de Kalven y Zeisel (1966), se sostiene que, pese a que no se pueda afirmar de modo definitivo, los datos apuntan a que la mayoría de las absoluciones no son aberraciones irracionales, es decir, grotescas concesiones hechas a gente culpable, sino veredictos con indicadores al menos consistentes con la inocencia material de los absueltos. La cuestión, afirma Findley, es que la cautela con la que Givelber y Farrell piden tomar sus resultados es aprovechada por Laudan para descartar su análisis, salvo por la mención de pasada de parte de los autores de que la dimensión del problema de las absoluciones falsas ha sido un misterio y quizá continúe siéndolo por un buen tiempo (Findley, 2018: 1276-1281).

Por su importancia, vale la pena detenernos un poco en el trabajo de Givelber y Farrell (2012). En él los autores analizan la información recabada en un estudio a gran escala realizado por el National Center for State Courts (NCSC) durante los años 2000 y 2001. Un objetivo prioritario de los autores fue determinar si los datos del estudio del NCSC eran o no, luego de transcurridas varias décadas, consistentes con la famosa «hipótesis de la liberación» proveniente del estudio de Kalven y Zeisel. De acuerdo con dicha hipótesis (Givelber y Farrell, 2012:. 5, 20-39), como se dijo antes, el jurado absuelve a las personas que los jueces habrían condenado porque se trata de casos en los que, ante sus ojos inexpertos en cuestiones probatorias, le surgen dudas de que el estándar PMATDR haya sido satisfecho. Estas dudas, que por la inexperiencia de sus miembros es comprensible que tenga, no obstante, carecen de fundamento pues, en realidad, las pruebas cumplen con dicho estándar (de lo cual los miembros del jurado se darían cuenta si tuvieran la capacidad que dan los años de no dejarse llevar por las apariencias). Pero una vez que estas dudas surgen (repito, injustificada, aunque comprensiblemente), el jurado renuncia a tener que seguir lidiando con las complejidades probatorias y, en su lugar, prefiere mejor aprovechar la oportunidad para mostrarse compasivo con el acusado y sus circunstancias particulares, y para canalizar, a través de la absolución de aquel, las críticas que pudiera tener contra el derecho penal sustantivo, por ejemplo, que quizá es injusto que criminalice la conducta imputada, o tal vez que la pena mínima que el acusado enfrentaría es ya, de por sí, muy severa, etc. A esta se le llama la *hipótesis de la liberación* porque es justo eso, una hipótesis, una conjetura. Pero no cualquiera, sino aquella en la que, en su gran mayoría, coincidieron los jueces consultados al responder a la pregunta de por qué creían que el jurado había absuelto en los casos en los que ellos habrían condenado. Y la liberación a la que se refiere no es solo la de los acusados respectivos, sino también la del propio sistema de justicia. ¿Liberación de qué? De la crítica de la que hubiera sido objeto si, al contrario, el estudio de Kalven y Zeisel hubiese concluido que la mayoría de las absoluciones tenían que ver con personas realmente inocentes, pues eso habría implicado que algo anómalo estaba ocurriendo con la percepción que la policía y las fiscalías tienen de las pruebas que terminan acopiando. La anomalía en cuestión consistiría en que ellos frecuentemente piensan que las pruebas son suficientes para satisfacer el estándar PMATDR cuando no lo son de verdad. Por su parte, la misma *hipótesis de la liberación* hace que el jurado salga igualmente bien parado, pues mostraría que este hace una gran y honorable contribución: atemperar los rigores de la ley, volver más humano al sistema. Esto último fue muy importante en la época del estudio de Kalven y Zeisel debido a que en

ese momento se cuestionaba seriamente si convenía o no mantener la institución del jurado, si aportaba algo y, en su caso, qué era específicamente. De modo que gracias a la *hipótesis de la liberación* el sistema en su conjunto resultó reivindicado al tiempo que la creencia de que la mayoría de las absoluciones que se pronuncian lo son de personas materialmente culpables se consolidó en la sociedad estadounidense.

Expuesto lo anterior, vayamos entonces a las conclusiones a las que llegan Givelber y Farrell (2012: 67-98, 137-144) en su análisis: De acuerdo con los autores, como ocurrió en el estudio de Kalven y Zeisel, los jueces siguen manifestando su desacuerdo con una proporción muy alta de los casos en los que los jurados absuelven a los acusados respectivos. Asimismo, luego de más de cinco décadas, la opinión de los jueces acerca de las razones que llevan al jurado a absolver (en esos casos en que ellos habrían condenado) sigue alineándose con la *hipótesis de la liberación* antes referida. Lo interesante es que, dado que el estudio del NCSC recabó también las opiniones de los miembros del jurado en los casos contemplados, se puede ahora contrastar ambas perspectivas. Estos sostienen que sus decisiones de absolver estuvieron de hecho basadas en la valoración de las pruebas, es decir, no en la actitud compasiva con los acusados y crítica del derecho penal por la que, según los jueces, ellos (los miembros del jurado) se decantan cuando enfrentan casos marginales, mismos que, de nuevo, de acuerdo con los jueces, son marginales solo en apariencia, de lo cual se daría cuenta un experimentado juzgador de cuestiones probatorias. Y claro, del hecho de que los miembros del jurado sostengan que sus decisiones de absolver están basadas en las pruebas practicadas en la audiencia respectiva, no se sigue necesariamente que su valoración de ellas sea correcta, o sea, no se sigue con definitividad que dichas pruebas genuinamente establezcan una duda razonable de la culpabilidad del acusado y, menos aún, que apoyen su inocencia material. En efecto, puede ser que, con todo, los jueces se desempeñen mejor como juzgadores de los hechos, que su larga experiencia como profesionales en este campo haga una gran diferencia, y que, por tanto, estén en lo correcto al sostener que cuando el jurado absuelve se termina liberando a alguien muy probablemente culpable. Para sopesar esa posibilidad, en este punto es relevante atender a las cuestiones probatorias que Givelber y Farrell identifican como los factores que influyen determinantemente en las absoluciones con las que los jueces estuvieron en desacuerdo; factores a los que la *hipótesis de la liberación* se refiere solo de modo genérico, considerándolos implícitamente, recuérdese, como la cortina de humo que confunde a los inexpertos miembros del jurado, como los detonantes de las dudas por las que no deberían dejarse asaltar, como las apariencias por las que no deberían dejarse llevar. Esos factores son cuatro: el primero es una defensa activa del acusado que, además de una historia alternativa plausible, haya presentado uno o más testigos que le aporten algún grado de corroboración a dicha historia. El segundo es que el propio acusado haya declarado en juicio. El tercero es que, del contrainterrogatorio del que haya sido objeto por parte de la fiscalía, no haya surgido que tiene antecedentes penales —cuya existencia, de ser el caso, seguramente habría sido señalada por aquella, pues es en estas circunstancias cuando la ley le permite aportar esa prueba sin mayor complicación—. Y el cuarto es que el acusado, al declarar, haya invocado su inocencia material y que por ello su caso ha llegado hasta la etapa de juicio oral —dando a entender así que a lo largo del procedimiento se ha mantenido incólume en su postura de no haber cometido el delito que se le imputa, incluso ante los tratos que en las etapas previas la fiscalía le haya ofrecido—. Pues bien, que precisamente esos factores sean a lo que los jueces se refieren como la cortina de humo que hay que saber disipar, como las apariencias en cuyo engaño no hay que caer, como aquello que la experiencia enseña a no ponerle atención, es indicativo, más que de una valoración imparcial de las pruebas, de una suerte de prejuicio institucionalizado en contra de los acusados. Esto parece verse reforzado por una reflexión que al respecto hacen Givelber y Farrell, y por un dato adicional que aportan, en mi opinión, también muy preocupante. La reflexión es que los jueces parecen ser vulnerables al enfoque que en la práctica genera mayores réditos a la policía y las fiscalías, a saber, el presumir la culpabilidad de los acusados una vez que hay (o parece haber) razones para sospechar de ellos. Este enfoque, por un lado, tiene el efecto de que la policía no investigue (o no a fondo al menos) la historia alternativa que el acusado pudo haber manifestado, y que se aboque a la recopilación casi exclusiva de pruebas de cargo; y, por otro, que la fiscalía termine imputando los cargos respectivos solo sobre la base de esa perspectiva parcial. Complementando lo anterior, los jueces que con cierta frecuencia desestiman

su parte, con respecto a los falsos positivos, la intuición de partida habría sido que son mucho menos frecuentes y costosos que la otra clase de veredictos erróneos. E igualmente, el autor parece haber sobrevalorado los elementos corroborativos disponibles y subvalorado los que pudieran refutarla. Finalmente, en lo que concierne a las propuestas para solucionar el problema de la cantidad supuestamente exorbitante de falsos negativos, parece que solo sus ventajas fueron tomadas en cuenta[27].

Si la anterior es una caracterización plausible, en el proceder de Laudan parece entonces haber rastros o manifestaciones de lo que técnicamente se conoce como un «sesgo de confirmación»[28]. ¿En qué consiste ese sesgo cognitivo? En un conjunto de estrategias que inconscientemente se despliegan para procurar que no se modifique, sino al contrario, para que solo se robustezca, al menos ante nuestros ojos, la opinión, perspectiva, preconcepción o impresión que ya se tiene con respecto a alguna cuestión. Entre las estrategias aludidas se encuentran precisamente la magnificación del papel y valor de la información que corrobora la posición preconcebida de que se trate y, por supuesto, la minimización (e incluso, anulación) del papel y valor de la información que pudiese ponerla en entredicho (Lidén, 2018: 53-55).

Pero, ¿cuál podría ser la posición prematura o preconcebida de la que partió Laudan? Si hubiese un sesgo confirmatorio esta debería poderse identificar. Sostengo que no es otra que la consideración del procedimiento acusatorio estadounidense, por parte del autor, como uno demasiado *amigable con la absolución*[29]. Esta consideración irrumpe en la escena desde su primer libro en este ámbito (Laudan, 2006[30]; 2013[31]). Y lo hace con la fuerza de una premonición del panorama sombrío sobre la frecuencia y costos de los falsos negativos que termina ofreciéndose en la obra posterior (Laudan, 2016).

---

los cargos en contra de los acusados son vistos con desaprobación; pero, además, su propia experiencia cotidiana en la ratificación de cientos y cientos de acuerdos entre los fiscales y los acusados puede ser interpretada como confirmatoria de la idea de que la mayoría de ellos son culpables, pues es natural que automáticamente se piense que eso, que son genuinamente culpables, es lo que explica que la policía y los fiscales respectivamente hayan recabado y presentado las pruebas de cargo de que se trate, y que los acusados admitan «voluntariamente» los hechos que se les imputa (2012: 116-118). Por su parte, el dato adicional que, en mi opinión, termina disminuyendo aún más la credibilidad de la *hipótesis de la liberación* y favoreciendo por tanto la hipótesis del prejuicio institucionalizado, tiene que ver con el componente racial en las opiniones de los jueces. Como destacan Givelber y Farrell, en síntesis, los jueces resultaron ser más propensos a condenar cuando los acusados son personas negras que cuando son hispanas o blancas. Asimismo, cuando las personas en cuestión eran negras, los jueces mostraron un mayor acuerdo con el veredicto del jurado que el que tendrían tratándose de personas blancas o hispanas, si este era condenatorio; y un menor acuerdo si el veredicto era absolutorio (2012: 120-136).

[27] Véanse las críticas 3 a 7 a Laudan que Aguilera (2025, p. 253-255) atribuye a Gardiner (2017), y las críticas 3 y 5 que el mismo autor (Aguilera, 2025: 257-261) atribuye a Findley (2018).

[28] Para conocer el estado del arte de la investigación sobre el sesgo de confirmación y particularmente cómo este se manifiesta en el ámbito procesal penal, véase (Lidén, 2020; 2018).

[29] A la cual me referí en una sección previa (la sección 2), llamándole *diagnóstico general*.

[30] Para su versión original en inglés.

[31] Para su versión traducida al castellano.

El punto es que ni siquiera la caracterización *preliminar* de aquel sistema, —repito, como uno demasiado amigable con la absolución—, parece haber sido el resultado de un balance neutral de los factores pertinentes. Sin embargo, que dicho diagnóstico fuese elaborado de un modo más imparcial era una expectativa fundada si se toman en cuenta las tareas que el propio Laudan (2006) atribuye a su fecundo e influyente proyecto de epistemología jurídica, cuyos cimientos quedan establecidos en la obra en cuestión. Veámoslo:

Como se sabe, a la manera de una extensión de lo que Laudan llama la «epistemología aplicada»[32], la novedosa disciplina de la epistemología jurídica busca 1) determinar cuáles de las reglas procesales penales vigentes en alguna región o sistema (incluidas las probatorias) promueven o facilitan la averiguación de la verdad[33], es decir, la minimización de los veredictos erróneos, tanto en el sentido de falsos positivos, como de falsos negativos, y; 2) proponer cambios en aquellas que se haya determinado que, al contrario, contribuyen a obstaculizar esos objetivos gemelos o, dicho de otro modo, a que aumente la probabilidad de que los referidos veredictos erróneos tengan lugar (Laudan, 2013: 23).

Planteadas así las cosas[34], parecería que lo que uno iba a encontrar en la obra seminal en comento sería un *esfuerzo simétrico* que, con el mismo ahínco, intentase detectar, para el caso de la obstaculización de la verdad, tanto reglas que incrementan el riesgo de que se cometan condenas falsas, como reglas que incrementan el riesgo de que se cometan absoluciones falsas.

Sin embargo, en su conjunto, el libro constituye, por una parte, una suerte de manifiesto en contra de lo que en él se identifica como una inclinación estructural desbordada del sistema acusatorio estadounidense hacia el pronunciamiento de absoluciones y, por la otra, una denuncia, en aquel entonces un tanto velada, del problema de la sobreabundancia de las absoluciones falsas al que esa exagerada inclinación conduce, que ya desde esta instancia se anticipaba que es grave.

Como muestra, considérense las extensas listas de lo que Laudan llama reglas probatorias y reglas procesales que obstaculizan la averiguación de la verdad (Laudan, 2013: 197-198, 204). Entre ellas, no hay ninguna que aumente el riesgo de condenar

---

[32] Que, siguiendo a Laudan (2013: 23), es el estudio orientado a determinar si las distintas formas de organizarnos para averiguar la verdad acerca de alguna cuestión o asunto preponderantemente empírico en el que estemos interesados —formas que dan lugar a sistemas sociales como, aunque no solo, las ciencias—, cuentan o no con un diseño adecuado (originalmente deliberado o no) a los efectos de permitirles la obtención de ese objetivo, el cual, dicho de un modo más directo, consiste en generar creencias verdaderas acerca del mundo (natural y/o social).

[33] De la verdad acerca de si un delito fue o no cometido y, en su caso, quién lo cometió.

[34] Obviando por ahora 1) el problema de que el acento del proyecto se pone en el análisis casi exclusivo de reglas (legislativas, pero, sobre todo, de origen jurisprudencial), en lugar de abrirse también a la consideración de lo que realmente ocurre en la práctica, que, para nuestra sorpresa, bien puede subvertir lo plasmado en aquellas, y; 2) el problema de su escasa e incluso nula atención a las fases previas a la del juicio oral propiamente dicha, en especial, a la etapa de investigación.

falsamente. Todas son, por volver a usar la expresión, amigables con la absolución (y, por tanto, factores que llevan al alza el riesgo de absolver falsamente)[35].

No es sino hasta diez años después[36] que Laudan (2016: 119-120) identifica algunas pocas de las reglas que aumentan el riesgo de una condena falsa[37]. Y no, no es que estas reglas no existieran desde antes; no es que de pronto hayan surgido en el periodo transcurrido entre su primer libro y el segundo. Es solo que, en esa primera incursión en la reflexión epistemológica sobre el proceso penal estadounidense, Laudan no consideró que ameritaran siquiera la mención de pasada que de ellas hace en su obra posterior.

Por otro lado, con respecto a la cuestión de las absoluciones falsas, ya desde esta primera obra puede constatarse, como adelanté, que para el autor constituyen un problema acuciante al decir que «[…] una de las tesis principales de este libro es que, dada la manera en que funcionan actualmente las cosas, *el sistema de impartición de justicia de los Estados Unidos permite, e incluso fomenta, la producción de más absoluciones falsas de las necesarias o deseables*» (énfasis añadido, Laudan, 2013: 111).

Asimismo, puede percibirse también ya el coqueteo del autor con la idea de la reincidencia cuando afirma que «[…] el costo principal de una absolución falsa es que un delincuente genuinamente culpable eluda el castigo correspondiente y quede libre, *quizá para cometer otros delitos*» (énfasis añadido, Laudan, 2013: 111). Y luego cuando sostiene que «[…] el mensaje que se envía a otros *delincuentes potenciales es que quizá también ellos puedan evitar ser perseguidos y castigados por los delitos que*

---

[35] Algunos ejemplos de estas reglas son: la que, aun si hubiesen sido voluntarias y estuviesen corroboradas, requiere la exclusión de las confesiones rendidas por los acusados si es que previamente no fueron informados de los derechos que, con base en el famoso caso Miranda, les asisten, o si fueron arrestados sin causa probable o fundada; la que, pese a su relevancia, prohíbe a la acusación tanto el uso de pruebas ilícitamente obtenidas, como el de las que pudiera haber encontrado siguiendo la pista de aquellas; la que impide a la acusación ofrecer pruebas de la personalidad y/o antecedentes del acusado, salvo que sea para impugnar las del mismo tipo que el propio acusado haya aportado a su favor; la que requiere de la inadmisión de pruebas que puedan prejuiciar injustamente al jurado en contra del acusado; la que impide informar al jurado de la negativa del acusado a responder los interrogatorios de la policía en las fases previas al juicio; la que impide que se apelen las absoluciones y que los miembros del jurado revelen lo discutido en sus deliberaciones, pese a que exista sospecha de graves irregularidades y/o de conducta inapropiada de parte de estos; la que no le permite a la acusación exigir que el acusado «descubra» sus pruebas, salvo que este se lo exija a ella primero; etcétera.

[36] Digo diez años después porque estoy considerando como referencia la versión original en inglés de ese primer libro, es decir, la de 2006.

[37] Como la que le otorga a los fiscales una amplia discreción con respecto a los cargos que puede imputarle al acusado, de modo que, en el contexto de la negociación respectiva estos pueden valerse de esa discreción para lograr que el acusado acepte al menos alguno de los cargos «a elegir», en principio, el menos lesivo de sus intereses (aun si no cometió el delito de que se trate); o la que les facilita a los fiscales la obtención de testimonios que incriminan al acusado permitiéndoles que ofrezcan tratos atractivos a sus supuestos cómplices, o bien inmunidad a otras personas respecto de los delitos que los fiscales pudieran achacarles (ahí la amenaza), a cambio de que renuncien a su derecho de invocar la quinta enmienda y, por tanto, de que declaren en contra del acusado en cuestión.

pudieran perpetrar. De hecho, presumiblemente *los índices de criminalidad podrían aumentar si se observa que las absoluciones falsas son un fenómeno frecuente*» (énfasis añadido, Laudan, 2013: 111).

Hablando de costos, es interesante percatarse de que los que en el trabajo en comento Laudan está dispuesto a atribuir a las condenas falsas, comprenden «[…] la mancha permanente en el buen nombre de la persona inocente que es erróneamente condenada, la privación de su libertad por el tiempo que dure el encarcelamiento, así como la pérdida de ciertos beneficios importantes asociados a su estatus de ciudadano (por ejemplo, en ocasiones, el derecho a votar)» (énfasis añadido, Laudan, 2013: 111). No es que haya habido aquí un gran esfuerzo por continuar dilucidando los costos de una condena falsa; lo que creo que vale la pena destacar, no obstante, es que al menos esos costos eran, en esta instancia, mayores que los que termina contemplando en su segundo libro (Laudan, 2016), los cuales, como se ha visto, se reducen al del inocente privado de su libertad (considerado por el autor tan víctima como la víctima del delito respectivo mismo), aunado a los delitos graves que en promedio al año cometerá el verdadero perpetrador del delito por el que al inocente equivocadamente se condenó.

En lo que toca a la frecuencia tanto de absoluciones como de absoluciones falsas, también desde aquí el autor muestra su sospecha de que es muy elevada. Contra quienes sostendrían que no se debe exagerar la dimensión de las segundas por pensar que la gran mayoría de las sentencias de los casos que llegan a juicio son condenatorias, Laudan (2013: 112) arremete primero citando la cifra de las 14 mil absoluciones que anualmente se pronuncian a nivel estatal y federal, la cifra de 300 mil acusados a quienes les son retirados los cargos, y el dato de que a la mitad de los sospechosos inicialmente arrestados (pero no acusados formalmente), se les deja ir. Ante esto, sostiene el autor, aunque seamos conservadores en la estimación de la cifra de culpables materiales que permitimos escapar de la justicia (en gran medida como resultado del elevado estándar de prueba imperante), podríamos estar hablando *de decenas de miles* de absueltos[38] falsamente por año (incluso de *centenas de miles*).

Un último asunto al que quiero llamar la atención es la tendencia de Laudan a emplear estudios empíricos, como el de Feeney y otros (1983), que explícitamente reconocen haberse basado en una sola de las perspectivas respecto de la probable culpabilidad o inocencia de los acusados: la perspectiva de la policía y de los fiscales. Estos estudios documentan que, en opinión de los actores mencionados, «… *la mayoría de los sospechosos que son arrestados, pero no condenados, son culpables*» (énfasis añadido, Laudan, 2013: 112). Laudan concede que la policía y los fiscales podrían estar equivocados, sin embargo, tiene grandes dudas de que sea así. Y la razón para dudar por parte del autor, la cual va en contra de la idea de que quienes rechazan el trato que la fiscalía pudo ofrecerles y optan por ir a juicio son en su mayoría gente inocente,

[38] En la acepción amplia de absolución, que comprende no solo las sentencias de esa modalidad, sino también las destimaciones de cargos.

es que no necesariamente es así. ¿Por qué? Porque, según Laudan (2013: 113-114), gracias al muy elevado estándar de prueba vigente, la cuestión de si se es realmente inocente o culpable pasa a un segundo plano en lo que concierne a la decisión de proseguir con un juicio. Lo importante y estratégico para tales efectos es, más bien, establecer si la fiscalía cuenta con pruebas de la entidad suficiente como para superar aquel estándar riguroso. Y cuando se juzga que no es así, eso es una buena noticia no solo para los acusados materialmente inocentes, sino también (y, sobre todo, según Laudan) para los materialmente culpables.

5. RÉQUIEM POR EL NÚCLEO DURO (O HARDCORE) DE LA EPISTEMOLOGÍA JURÍDICA

Aunado a lo dicho hasta el momento y yendo un paso todavía más atrás, considérese también la propia estructura del trabajo bajo análisis como indicio de una visión que ya venía sesgada, veámoslo: Luego del capítulo introductorio[39], vienen las dos grandes partes que lo conforman. En la primera (que va del capítulo II al IV), se abordan las cuestiones que atañen a lo que Laudan (2013: 173-174) denomina el «núcleo blando» (o *softcore*) de la epistemología jurídica, o sea, los dispositivos procesales que piensa que, *de forma exclusiva*, deben usarse para contribuir a que los errores previsibles se distribuyan de un modo socialmente aceptable, es decir, de un modo *que refleje sus costos relativos*. Por esta razón, en su conjunto, los dispositivos aludidos, —y repito, solo ellos, según Laudan—, deben sentar las bases o crear el entorno propicio para volver más probable que, cuando se cometan errores[40], estos preponderantemente pertenezcan a la clase de los menos costosos, los cuales, al menos en la mayoría de las sociedades occidentales (y por el momento), no son otros que las absoluciones falsas[41]. Esos dispositivos se corresponden con lo que antes llamé la tradicional estructura acusatoria tripartita conformada por la presunción de inocencia, la carga del Estado de probar la culpabilidad y un elevado estándar de prueba, aunque es a este último al que Laudan le presta la mayor atención[42].

Como puede observarse, entonces no es que Laudan rechace toda inclinación estructural hacia la absolución por parte del sistema acusatorio, lo cual implica que tampoco rechaza lo que esa inclinación conlleva, a saber, el aumento del riesgo (o de la probabilidad) de que se cometan más absoluciones falsas que condenas falsas (cuando el sistema incurre en la emisión de veredictos erróneos). Las cuestiones son cómo y cuánto aumentar esa probabilidad.

---

[39] Una magistral pieza analítica, por cierto.

[40] Posibilidad que, por más esfuerzos que hagamos, solo puede reducirse, más no erradicarse del panorama.

[41] Nótese que no se dice que estas no tengan costo alguno. Este puede ser muy alto; incluso tan alto como Laudan piensa. Lo que sí se afirma (por Laudan y la mayoría de autores) es que, comparada con una condena falsa, absolver al culpable es algo menos grave.

[42] Particularmente en los capítulos II y III.

La respuesta al cómo es únicamente mediante el empleo del estándar de prueba. Ello presupone la previa instauración de la presunción de inocencia, que le traslada la carga de probar la culpabilidad a la fiscalía, de modo que el estándar de prueba especifica el umbral de cuya satisfacción depende el desahogo exitoso de dicha carga.

Por su parte, la respuesta al cuánto tiene que ver, para Laudan, con el *adecuado empleo* de la teoría de las utilidades esperadas. En este orden de ideas, desde la perspectiva de quien diseña el sistema, el punto en el espectro de los grados progresivos de la culpabilidad aparente (de un acusado)[43] en el que se debería fijar el estándar de prueba[44], sería aquel en el que las utilidades de las cuatro modalidades posibles de un veredicto convergen[45].

A esta pauta, Laudan la denomina «principio de indiferencia». ¿Por qué? Porque justo en el punto en el que seríamos indiferentes con respecto a si debería absolverse o condenarse —o sea el punto en el que las utilidades de las cuatro modalidades de un veredicto coinciden o en el que las líneas que las representan se entrecruzan—, es donde el estándar probatorio debe establecerse. Ese punto en el espectro de valores progresivos y probabilísticos de culpabilidad aparente recoge toda la dosis de protección al acusado (en especial al materialmente inocente) que esperamos inyectar al sistema, es decir, toda la cuota de beneficio de la duda[46] que desearíamos racionalmente concederle (siempre y cuando la teoría de las utilidades esperadas haya sido correctamente aplicada, claro está).

Siendo así, y para no alterar el equilibrio alcanzado a través de este procedimiento, Laudan proscribe la concesión indirecta de más y más dosis de beneficio de la duda al acusado[47]. De ahí su férrea oposición, y el calificar como demasiado amiga-

---

[43] O sea, de la probabilidad de culpabilidad del acusado que resulta de la adecuada valoración de las pruebas disponibles.

[44] Espectro al que podríamos imaginar como una línea horizontal a la manera del eje X de un plano cartesiano.

[45] Dichas utilidades serían mayores o tendrían valores altos para el caso de las modalidades correctas (condena verdadera y absolución verdadera), pues representan sus beneficios, y serían menores o tendrían valores bajos para el caso de las modalidades erróneas (condena falsa y absolución falsa), pues representan sus costos.

[46] Beneficio de la duda en el sentido de la cantidad y calidad de las pruebas, por encima del punto que representa el valor 0,5 de culpabilidad aparente (es decir, por encima del estándar de la preponderancia de las pruebas), que, como sociedad estamos dispuestos a que el juzgador obvie o ignore en contra del acusado. O sea, toda la gama de valores de culpabilidad aparente, —repito, por encima del punto que representa el valor 0,5—, que estamos dispuestos a que ameriten la absolución del acusado (en efecto, pese a que en su contra existan pruebas en cuya base se pueda decir que este es probablemente culpable en términos materiales; de lo contrario, no tendría mucho sentido llamarle «beneficio de la duda»).

[47] Indirecta porque el único medio directo y legítimo que, según Laudan, debería emplearse para la concesión de la dosis racionalmente atemperada de beneficio de la duda al acusado consiste, como ya sabemos, en fijar el grado de exigencia que el estándar de prueba deberá tener para luego ser empleado por los juzgadores como criterio de suficiencia, o sea, como pauta para evaluar si la fortaleza del caso que la fiscalía presenta alcanza para poder dar válidamente por probada la hipótesis de culpabilidad cuya verdad defiende.

bles con la absolución o excesivamente garantistas, a las reglas que tienen el efecto de que se excluyan pruebas de cargo relevantes, y a cualquier otra que de algún modo le otorgue mayores ventajas al acusado. –«Siguiendo el procedimiento establecido, es decir, respetando el principio de indiferencia, ya se ha hecho todo lo razonable para canalizar la preocupación de no condenar erróneamente; el resto son concesiones gratuitas, injustificadas del todo, que terminan beneficiando sobremanera al acusado materialmente culpable, facilitando su absolución»–, sostendría el autor.

Establecida entonces la importancia capital del principio de indiferencia[48], la segunda parte del libro (conformada por los capítulos del V al IX) se aboca de lleno a intentar aplacar la rebelión que contra él se gesta por parte, como se dijo, de todas esas reglas que engrosan de más el cinturón de protección para el acusado que, según Laudan, previamente ya fue creado por virtud del estándar probatorio, superior al de la preponderancia de las pruebas, que resulte racional implementar[49].

Esta agenda, en mi opinión, es en realidad una continuación de lo que el autor definió como el *softcore* de la epistemología jurídica. ¿Por qué? Porque se centra en denunciar y sugerir que se modifique todo aquello que no sea consistente con el principio de indiferencia, es decir, todo lo que aumente la probabilidad de que los errores se *distribuyan* de modo aún más favorable a la absolución falsa de lo que ya de por sí ese principio garantiza.

El punto es que una agenda con esas características se queda corta respecto de lo que el capítulo V anuncia que será abordado en el resto del libro, a saber, supuestamente ahora sí las cuestiones del «núcleo duro» (o *hardcore*) de la epistemología jurídica. Esas cuestiones en teoría tienen que ver no ya con la contribución del sistema a que los veredictos erróneos se *distribuyan* del modo socialmente aceptable, sino con lo que de entrada haría *que fuesen menos* los errores que luego habrían de distribuirse de ese modo, es decir, con el perfil de las reglas que contribuirían a *minimizar* la ocurrencia tanto de falsos positivos, como de falsos negativos.

Sin embargo, esas cuestiones no son tratadas, o no con la neutralidad esperable. ¿Por qué? Porque, como he intentado mostrar antes, en el elenco de reglas que Laudan identifica y discute no hay reglas amigables con las condenas falsas, sino solo reglas amigables con los falsos negativos. Y, como también antes destaqué, no es que

---

[48] Principio que, vuelvo a decirlo, prescribe introducir en el sistema la dosis de beneficio de la duda que se espera que se conceda al acusado a través (y solo a través) del grado de severidad del estándar de prueba que resulte amparado por el uso correcto de la teoría de las utilidades esperadas.

[49] Entre las reglas que son discutidas a fondo están la que confiere a los acusados el derecho a guardar silencio, la que le permite a los testigos ampararse en la quinta enmienda para no autoincriminarse, las reglas que crean distintas categorías de testigos «privilegiados» (capítulo VI), la que ordena la exclusión de las confesiones obtenidas ilícitamente, así como de las pruebas que resulten de usar la información contenida en dichas confesiones (capítulo VII), el principio *non bis in ídem* que se invoca para impedir la apelación de las absoluciones (capítulo VIII), la regla que promueve la exclusión de las pruebas que injustamente puedan prejuiciar al jurado en contra del acusado, y la que prohíbe el empleo de pruebas obtenidas sin la respectiva orden de registro y cateo (capítulo IX).

no las hubiese, sino que el autor opta por no incluirlas o, quizá más bien, escapan de su foco de atención por lo arraigado que, a esas alturas, ya parece encontrarse el prediagnóstico de que lo que se tiene enfrente es un cuadro grave de hipergarantismo, que debe ser contenido de manera prioritaria y urgente.

6. LA DILUCIÓN DE LA ESTRUCTURA ACUSATORIA TRIPARTITA

En la segunda sección del trabajo, al referirme al que poco a poco ha emergido como el acérrimo enemigo de Laudan, el hipergarantismo, di a entender que consiste en el efecto que en la operación de los sistemas acusatorios tiene la actitud de insistir en la reducción por todas las vías posibles del riesgo de una condena falsa; actitud que suele traducirse en ir «más allá» de la dinámica probatoria a la que la tradicional estructura tripartita de estos sistemas los condiciona. Y sostuve también que dicha estructura está conformada por la presunción de inocencia, la carga del Estado de probar la culpabilidad y el estándar de «prueba más allá de toda duda razonable» (PMATDR).

Lo que pondré ahora de relieve es que ni siquiera esa estructura básica y en buena medida identitaria de los sistemas acusatorios queda a salvo del asedio al que Laudan los somete motivado por su afán de reducir la sobreabundancia de absoluciones falsas[50].

De lo anterior podemos percatarnos, por un lado, cuando entendemos que el principio de indiferencia por el que aboga Laudan[51] no es sino la extensión de la teoría de las utilidades esperadas al terreno de la decisión sobre el grado de exigencia del estándar de prueba penal. Y, por otro, al caer en la cuenta de que, al serlo, ese grado de exigencia ha quedado sujeto precisamente a aquello de lo que esta teoría toma su nombre: a las utilidades o costos relativos que estemos dispuestos a asignar tanto a los falsos positivos como a los falsos negativos. Dado que nada garantiza que dichas utilidades o costos, incluso suponiendo que han sido razonablemente dilucidados, permanezcan estables, igualmente fluctuante se torna la naturaleza del estándar de prueba penal.

A esa fluctuación Laudan (2013: 21-22) le da la bienvenida inclusive desde las primeras líneas de la obra en comento cuando, para abrir, describe los tres grandes objetivos que según él son *evidentes* si se observa de cerca el sistema de justicia penal estadounidense (y el de la mayoría de los países occidentales). Me referiré a continuación, a los dos más importantes: el primero de ellos es el de averiguar la verdad sobre el delito en cuestión, por tanto, el de evitar el pronunciamiento de veredictos falsos. El segundo, parte de la sensata advertencia de que, por más esfuerzos que hagamos,

[50] Problema de cuya gravedad el autor pretende habernos persuadido a la luz de la supuesta prueba empírica que aporta en su segundo libro.

[51] Véase la primera parte de la sección previa.

una cuota de esos errores será inevitable, tanto porque el acervo de pruebas del que se logra disponer sufre regularmente de algún grado de incompletitud, como porque el juicio humano, sin importar lo bienintencionado y competente que pudiera ser, es falible. Siendo así, ese segundo objetivo consiste en dilucidar cuál de las clases posibles de error es la más costosa, una condena falsa o una absolución falsa; para luego contribuir a su reducción en detrimento de la otra. Nótese que este objetivo no se formula directamente como el de contribuir a la reducción de las condenas falsas pagando por ello el precio del crecimiento de las absoluciones falsas, sino que su articulación se mantiene neutral, es decir, sin asumir de forma clara ese compromiso. En mi opinión, ya desde aquí se está preparando el terreno para posteriormente presentar al estándar de prueba como un dispositivo maleable o ajustable, sujeto a los designios de la teoría de las utilidades esperadas.

Y claro, esa maleabilidad podría dar pie a que el grado de exigencia del estándar se eleve por encima del valor probabilístico con el que suele asociarse a la expresión «prueba más allá de toda duda razonable» (PMATDR), que es el valor de 0.9; pero también a que caiga por debajo de dicho valor. Sin embargo, que el estándar solo vaya en descenso es de lo que Laudan quiere asegurarse al plantear las cuestiones a las que me referiré a continuación.

Pero antes, cabe mencionar que si la asignación de costos, en esa espiral de descenso en la que podría caer, se desplomase hasta llegar al punto de que ambas clases de veredictos erróneos se consideren igualmente costosas, no tendría sentido entonces preservar la presunción de inocencia, ni la carga del Estado de probar la culpabilidad, pues en tal caso, el estándar que correspondería emplear sería el de la preponderancia de las pruebas (es decir, el de 0,5+), o incluso uno menor si las cosas llegan al extremo de que sean las absoluciones falsas, en lugar de las condenas falsas, el error percibido como el más costoso [52].

Ahora sí, vayamos a las cuestiones que el autor pone en la mesa para persuadirnos de que el estándar racional debe ser menos exigente que PMATDR:

1) La expresión «PMATDR», de inicio, está afectada de una dosis muy considerable de indeterminación. Ni siquiera las mentes jurídicas más sagaces han podido dar cuenta de su significado (Laudan, 2013: 66-68). Al contrario, han llevado las cosas al punto de sostener que se trata de una noción no solo autoevidente, sino elusiva de cualquier intento de definición (pp. 83-88). Esto no sin antes haber propuesto supuestas clarificaciones que terminan yendo en contra de la certeza jurídica al truncar la aplicación uniforme del estándar en cuestión (pp. 68-74). ¿Por qué? Porque suelen hacer que la satisfacción de PMATDR dependa de cuestiones muy subjetivas, por ejemplo, de si el juzgador considera su convencimiento de que el acusado es culpable lo suficientemente parecido en grado, al convencimiento que, en otros contextos de

[52] Escenario para el cual es más apropiada una presunción de culpabilidad (en lugar de una de inocencia), cuya condición de derrotabilidad sea la prueba, a un estándar muy elevado, de que el acusado no cometió el delito de que se trate. Las cosas al revés, para decirlo rápido.

su vida cotidiana, ha experimentado como paso previo a la toma de decisiones prácticas importantes (pp. 68-69).

2) El estándar PMATDR y la ratio de 10 absoluciones falsas por cada condena falsa que (supuestamente) contribuye a implementar, pueden resultar muy inapropiados en virtud de un escenario imaginable en el que, en un determinado periodo, haya habido digamos 100 veredictos y 9 de ellos fuesen condenas falsas. Si la distribución del error tiene que ser de 10 a 1, debería por tanto haber 90 absoluciones falsas, con lo cual se llega al absurdo de que solo un veredicto podría ser materialmente correcto (ya sea de absolución o de condena), a saber, el veredicto restante para completar los 100 que se ha supuesto que se emitieron (pp. 116-117).

3) La ratio de 10/1 tan comúnmente asociada a PMATDR parece más una intuición acerca de la distribución de errores deseada (bienintencionada, sin duda), que una ratio bien fundada en el empleo correcto de la teoría de las utilidades esperadas. Al ser atemperados los costos de las condenas falsas y de las absoluciones falsas mediante la consideración de las utilidades o beneficios de las absoluciones verdaderas y de las condenas verdaderas, —que es a lo que, según Laudan y Saunders (2009) conduce la aplicación depurada de la teoría en cuestión—, no parece haber forma racional de obtener esa distribución, conocida como ratio Blackstone, sino solo distribuciones menos desproporcionadas en lo que atañe a la manera en que favorecen a las absoluciones falsas. De hecho, como se vio en el tercer apartado del trabajo, según los cálculos de Laudan resultantes de tomar en cuenta su métrica de víctimas y los «datos» en los que se basó, las condenas falsas son solo 2 (y no 10) veces más costosas que las absoluciones falsas.

4) Algo que puede estar causando la distorsión de costos implícita en la ratio Blackstone es la automática asociación que el ciudadano común (y aparentemente también el legislador) suele hacer entre el derecho penal y la expresión más severa de la potestad punitiva conferida al Estado: la pena de muerte (pp. 96-101). PMATDR podría ser el estándar apropiado para los delitos que enfrenten esa consecuencia, y *solo para esos casos*, pues, por definición, los daños a la víctima de una condena falsa en esas circunstancias, son irreparables; pero esta medida, la pena de muerte, no es, ni de cerca, de aplicación generalizada en los Estados Unidos (ni en buena parte del mundo). Es más, hay una buena cantidad de delitos que ni siquiera ameritan estancias en prisión como castigo. Por tanto, no parece razonable tampoco en esos casos que PMATDR sea el estándar aplicable, pues muestran que la consecuencia de un veredicto condenatorio en materia penal a veces no se compara incluso ni con ciertas repercusiones de perder un asunto en el ámbito civil, que, por ejemplo, pueden tener que ver con el pago de una suma de dinero muy cuantiosa a título de daños y perjuicios. Y cabe recordar, nos pide Laudan, que en el ámbito de la mayoría de dichas controversias civiles ya de entrada está vigente el estándar de la preponderancia de las pruebas, es decir, uno bastante inferior a PMATDR, el cual meramente requiere, para que una afirmación fáctica sea declarada probada, que esta sea más probable que su negación (pp. 93-95).

5) La estructura acusatoria tripartita ya de hecho se trastoca en Estados Unidos (y quizá en muchas regiones más). Esto ocurre cuando la normativa procesal traslada al acusado la carga de probar, usualmente al estándar de la preponderancia de las pruebas, las llamadas defensas afirmativas (como la inimputabilidad o la defensa propia) que haya invocado, lo cual implica que en el acusado cae la losa de tener que demostrar activamente su inocencia (pp. 165-170).

7. EN NOMBRE DE LA CIENCIA, LA FILOSOFÍA Y LA EVIDENCIA EMPÍRICA

Interesante es observar cómo a esta cruzada contra el hipergarantismo —que, como se ha visto, no se limita a identificar y sugerir el cambio de las reglas amigables con la absolución adicionales a la tradicional estructura acusatoria tripartita, sino que también atenta contra ella (especialmente contra el estándar PMATDR)—, Laudan la presenta como aquello que resulta de asumir una perspectiva neutral, filosófica, científica y sensible a la consideración de la información empírica pertinente; en suma, como algo que no podemos sino aceptar si hemos de mantenernos dentro del perímetro de la racionalidad.

En respaldo de lo anterior, considérese el experimento mental que Laudan (2013: 25-32) se propone hacer con respecto a un procedimiento penal: suponer cómo luciría, qué configuración tendría, si la averiguación de la verdad sobre si un delito ocurrió y si fue así quién lo cometió, fuese nuestra única preocupación.

Dicha forma de proceder es comparada por el autor (pp. 27-29) con la estrategia que Galileo siguió para comprender más nítidamente el fenómeno de la caída de los cuerpos, la cual consistió en suprimir momentáneamente la consideración de factores obviamente relevantes (como la forma y el peso de aquellos, o el medio por el que caen), para centrarse solo en el papel de la fuerza de gravedad. De ahí el emparentamiento de su metodología con la ciencia[53] y, por supuesto, con la filosofía, misma que con frecuencia recurre justamente a experimentos mentales como este para probar algún punto y, sobre todo la de corte más analítico, a presentar de la forma más clara y precisa posible el aparato conceptual del que se valdrá; tal como también hace Laudan al plantear las distinciones que para él serán cruciales, como la que diferencia entre la culpabilidad y la inocencia materiales y la culpabilidad y la inocencia probatorias (pp. 36-37, 147-151), o la distinción entre diversas clases de error (pp. 33-44): error lógico o epistémico, error procesal, error inferencial, etcétera.

Luego de explicar las características de su experimento mental, Laudan (*Ibidem*, p. 31) también manifiesta su inclinación por la *naturalización* de la filosofía. Y de ahí su vínculo con el acopio de evidencia empírica siempre que esté disponible y sea

[53] Por supuesto, en adición a ser él, Larry Laudan, uno de los más prestigiosos filósofos de la ciencia de su época, quien la propone.

relevante para la cuestión filosófica que esté siendo abordada. Esto supuestamente se muestra en todo su esplendor en el diagnóstico específico con respecto al panorama de las absoluciones falsas que el autor ofrece en su segundo libro (Laudan, 2016).

Continuando con el experimento mental aludido, Laudan plantea el objetivo del mismo afirmando lo siguiente: «[…] si hemos de *comprender adecuadamente el núcleo de las cuestiones epistémicas que están en juego en el diseño y operación de los procesos penales*, es mejor que, en un principio y *solo temporalmente*, hagamos a un lado cualquier otra consideración» (énfasis añadido, Laudan, 2013: 27).

¿A qué otras consideraciones se refiere principalmente el autor cuando sugiere que es mejor que las hagamos a un lado de forma momentánea? Justamente a las que tienen que ver con el interés en brindar al acusado protecciones en contra de una condena falsa. Esto puede constatarse en pasajes como los siguientes:

> *1)* «[…] mi intención … consiste en abrir un espacio conceptual en el que podamos discutir cándidamente […] *sin toparnos inmediatamente con la supuestamente infranqueable objeción del tipo «pero x es un derecho» o «x está ordenado (o prohibido) por la Constitución»* […] trataré (por tanto) de adherirme al punto de vista de que *cuanto menos se hable de derechos, tradiciones y doctrinas jurídicas, y de derecho constitucional, será mejor*» (énfasis añadido, Laudan, 2013: 28-29).
>
> *2)* «[…] en tanto los juristas sigan creyendo […] que ciertas reglas procesales establecidas judicialmente, *como, por ejemplo, la supresión de confesiones obtenidas mediante coerción*, facilitan y promueven la averiguación de la verdad, —cuando, de hecho, tienen el efecto opuesto—, *no puede haber más que confusión en lo que concierne a la cuestión de si se está contribuyendo a la verdad y cuándo*» (énfasis añadido: 31).
>
> *3)* «[…] me aproximo a las cuestiones referidas *como un filósofo que observa el fenómeno jurídico desde fuera*, y no como un abogado que opera desde dentro […] desde esta óptica *me encuentro menos interesado en cosas como los derechos de los acusados (en contraste con el interés que pudiera tener en estos asuntos un promotor de las libertades civiles o un abogado defensor),* y más preocupado en cuán efectivo es el sistema de justicia penal en términos de producir fallos verdaderos» (énfasis añadido: 31-32).

Ahora bien, que la supresión de consideraciones ajenas al objetivo de averiguar la verdad dure solo lo que tome la realización del experimento mental mencionado, del que supuestamente se extraerán lecciones importantes sobre los rasgos de una genuina configuración promotora de la verdad del procedimiento penal, parece una buena noticia en virtud de la promesa de ser retomadas en una segunda fase.

El problema es que, una vez identificadas las lecciones a las que conduce su experimento mental, estas son ahora presentadas como la única ruta imparcial y racional que se puede seguir, o sea, como un conjunto de principios prístinos, limpios de impurezas al haber sido expuestos al fuego de la razón desinteresada. De ahí la disposición de Laudan (2013: 28) a no abandonarlos.

¿Por qué esto es un problema? Porque con base en el recorrido que nos ha traído hasta aquí, es plausible afirmar que a esos principios no se llegó de forma tan desinteresada. Esto parece especialmente claro en el caso del principio de indiferencia, pues con su sola postulación comienza ya la embestida de Laudan contra la inclinación estructural del proceso acusatorio hacia la absolución, con la que, en teoría, está de

acuerdo. Como sabemos, dado su vínculo con la teoría de las utilidades esperadas, dicho principio vuelve al grado de exigencia del estándar de prueba dependiente del veredicto al que conduzca el empleo de esa teoría (que a su vez depende de las fluctuantes utilidades o costos que les sean asignados a ambas clases de veredictos erróneos); y como sabemos también, dados los argumentos del autor, especialmente el que tiene que ver con el empleo correcto de la teoría en cuestión, ese grado de exigencia no puede racionalmente corresponderse con PMATDR, sino con alguno inferior, aunque en principio no tan bajo como el de la preponderancia de las pruebas.

En este orden de ideas, es cierto que las consideraciones ajenas a la averiguación de la verdad son eventualmente retomadas (a las que Laudan considera ligadas a la defensa de valores distributivos del error y extra-epistémicos). La cuestión es que vuelven teniéndose ya el telón de fondo conformado por el principio de indiferencia, cuya observancia, recordemos, según Laudan, garantiza que toda la dosis de beneficio de la duda a favor del acusado que es racional inyectar al sistema, le ha sido ya concedida.

De modo que sí, como vengo diciendo, esas consideraciones reaparecen, pero lo hacen revestidas de un halo de ilegitimidad o, al menos, de sospecha, toda vez que Laudan detecta que son invocadas como el fundamento de las reglas respecto de las cuales ya está dispuesto 1) a tildar de excesivamente amigables con la absolución (pues otorgan ventajas al acusado adicionales a la supuesta gran concesión que a su favor se ha hecho elevando al estándar de prueba aplicable por encima del de la preponderancia de las pruebas) y, por tanto, 2) a sugerir su eliminación.

Por su parte, como es de suponerse, algunas de esas reglas son profundamente apreciadas, y creo que, con mucha razón, pues le confieren al acusado derechos tan importantes como el de no cooperar con la policía en la etapa de investigación, el de guardar silencio durante su juicio, o el de que las pruebas en su contra, incluida su confesión, sean excluidas si fueron ilícitamente obtenidas. De ahí la resistencia, proveniente ahora de la contraparte jurídica de la ecuación, a que simplemente desaparezcan.

8. UNA PECULIAR TEORÍA UTILITARISTA DE LA JUSTIFICACIÓN DEL CASTIGO

Pienso que es plausible sostener que el origen de la cruzada de Laudan contra el espectro de las absoluciones falsas tiene que ver con la suscripción por parte del autor de una específica teoría utilitarista acerca de la justificación del castigo. Antes de esbozar dicha teoría voy a referirme a un conjunto de premisas con las que esta parece más compatible y que, en virtud de ello, le sirven de punto de partida. Estas son: 1) la criminalidad es un fenómeno persistente y en ascenso que desborda las capacidades limitadas del sistema de administración de justicia penal; 2) una gran proporción de quienes cometen delitos, sobre todo delitos graves, son o se convierten en reincidentes, en buena medida debido a un sistema inefectivo que no actúa con la fuerza represiva suficiente para ponerle un alto a sus carreras criminales, al menos

durante un tiempo; 3) la respuesta insuficiente del sistema se agrava en virtud de una serie de concesiones a favor de los acusados, fundadas más en el apego acrítico a ciertas tradiciones y mitos heredados, que en su racionalidad; 4) la perspectiva de quienes oficialmente combaten este flagelo, —policías, fiscales e incluso los jueces—, es la más fiable, especialmente respecto del grado en que las pruebas recabadas en contra de los acusados justifica la inferencia de que son culpables, y; 5) es *urgente* que se tomen medidas al respecto.

Con base en estas premisas[54], la teoría acerca de la justificación del castigo que pretendo poner de relieve respondería a la pregunta de por qué o de qué sirve castigar, del modo siguiente: porque imponiéndolo, especialmente en la modalidad de privar de la libertad al condenado, se da una respuesta proporcional a lo alarmante del problema de la criminalidad, contribuyendo a contenerla e idealmente a reducir sus índices; en el mejor de los casos, disuadiendo a los criminales de que cometan delitos en el futuro mediante el estímulo desagradable que representa el estar en prisión y, en el peor, cuando estos son ya incorregibles, incapacitándolos para que continúen dañando a otros, por un tiempo al menos (de preferencia prolongado).

Desde una perspectiva así, es natural que la inclinación estructural del sistema acusatorio a absolver sea vista, primero, no como un dique infranqueable que debe implementarse para limitar los abusos en los que el Estado puede incurrir, sino como un acto de magnanimidad; y, segundo (a consecuencia de lo anterior), como una política que debe ser constantemente revisada, teniéndose siempre en mente el costo en víctimas atribuible a las absoluciones falsas. Pero nada garantiza que esa inclinación deba mantenerse, pues estaría justificado eliminarla si el costo de las absoluciones falsas se llegara a equiparar con, o llegara a superar, el de las condenas falsas.

9. REFLEXIÓN DE CIERRE: ESBOZO DE UNA TEORÍA ALTERNATIVA DE LA JUSTIFICACIÓN DEL CASTIGO Y DE UN PROGRAMA DE EPISTEMOLOGÍA JURÍDICA BASADO EN ELLA

La teoría sobre la justificación del castigo que en este apartado esbozaré como alternativa a la que previamente sostuve que es plausible considerar como el fundamento del proyecto de epistemología jurídica de Laudan[55], es una teoría 1) que predica del castigo la función de entablar con el condenado un diálogo, una comunicación, de carácter moral[56], y 2) que prioriza dicha función[57]. Por tanto, a dife-

[54] Cuyo listado no ha pretendido ser exhaustivo.

[55] Y, en ese sentido, como el origen de su cruzada contra el espectro de las absoluciones falsas.

[56] Sigo de cerca en este aspecto a Lippke, 2024; 2016; 2011, quien a su vez se nutre de las ideas, entre otros, de Duff, 2001; 1996; 1986 y Hampton, 1992.

[57] Nótese que se habla de priorizar la función referida y no de considerarla como la única posible o la única importante, lo cual abre un espacio para la persecución simultánea de objetivos como retribuir

rencia de la que atribuyo a Laudan, la teoría que me encuentro presentando, apenas rudimentariamente, parte de suponer que la persona condenada es un agente moral, y no (o no únicamente) un ente autointeresado que solo entiende de reforzamientos positivos y negativos de su conducta, o una fuerza del caos con la que no se puede razonar, sino solo contener coactivamente para evitar que siga dañando a la sociedad.

Esta teoría supone también que, al castigarse a la persona condenada, especialmente privándola de su libertad, se le expresa una censura o un reproche preponderantemente moral, cuya comprensión como tal se espera de dicha persona [58]. De ella se espera, además, que entienda las razones de ese reproche (es decir, que se trata de un reproche fundado) y de que este tenga que llevarse a cabo de manera pública [59].

Ahora bien, para para que ese reproche a la persona condenada sea legítimo, el Estado primero tendría que revestirse de la autoridad moral para hacerlo; tendría que ganarse el derecho de ser el interlocutor moral que, de acuerdo con la teoría en comento, pretende ser al castigar. El itinerario para ganarse ese derecho es, en efecto, complicado y en muchos aspectos seguramente controversial. No obstante, en mi opinión, algunos de los rubros generales que lo conforman son relativamente claros, los cuales, por supuesto tienen que ver con el procedimiento que sirve de antesala del reproche, pero incluso con cuestiones previas a este y también ulteriores [60].

---

al perpetrador por sus actos, controlar o reducir la criminalidad, desincentivar o hasta incapacitar al sentenciado, rehabilitarlo, etcétera.

[58] Y no que conciba a la amenaza y/o imposición de castigos como una mera razón prudencial a incluir en sus deliberaciones futuras, o como una medida que físicamente, aquí y ahora, le impide continuar entregándose a sus impulsos predatorios.

[59] De lo primero, una razón es que, dado su estatus como agente moral, se esperaba de la persona condenada que se hubiera percatado de las consideraciones morales que respaldan la prohibición que, sin justificación ni excusa, negligente o dolosamente, quebrantó; y, asimismo, que se hubiera dejado guiar en su deliberación práctica por el peso de esas consideraciones. Y otra es que dicho reproche, al traducirse en la privación de su libertad, le da a la persona condenada la oportunidad de aprovechar ese periodo para transitar un ciclo que idealmente la lleve al arrepentimiento y a intentar reparar el tejido social lacerado por su conducta, pasando por el remordimiento de consciencia y por el duelo resultante de haber perdido la oportunidad de no cosificar a su víctima y de haberla tratado con el respeto que su condición como fin en sí misma requería. De lo segundo, es decir, de que el reproche tenga que llevarse a cabo públicamente, la razón es que el castigo constituye, a su vez, un mensaje que se manda al resto de la sociedad en el sentido de que sus valores fundamentales siguen y continuarán siendo observados, y de que, por tanto, tenemos razón para seguir aportando a la convivencia pacífica. Pero el castigo también constituye un mensaje públicamente enviado a la víctima y/o sus allegados, a quienes les comunica que la sociedad, por vía de sus autoridades, se solidariza con ellas, que empatiza con su sufrimiento y que se encuentra haciendo lo que razonablemente está a su alcance para reafirmar su valor como personas, para al menos evitar que la persona condenada vuelva a desviarse del ordenamiento, y para que repare en parte el daño sufrido.

[60] Entre las previas cabe destacar las condiciones socioeconómicas imperantes en la sociedad y la labor de tipificación de conductas llevada a cabo por el legislador; y entre las ulteriores, el entorno en el que permanecen privadas de su libertad las personas condenadas. En los aludidos ámbitos socioeconómico y penitenciario, el Estado debería esforzarse por imponer a las personas unas condiciones aceptables en el sentido de mostrar un mínimo de consideración y respeto a su dignidad. Y en ambos casos también, esa consideración y respeto deberían traducirse en el reconocimiento y consolidación de

En lo que respecta al procedimiento[61], este se vería en retrospectiva como una extensión de la comunicación moral que se entabla con la persona que resultó condenada, es decir, cuando esta tiene apenas el estatus de acusada, e incluso antes, cuando solo es sospechosa de estar involucrada en el delito que se investiga. En este orden de ideas, los rasgos de ese procedimiento deben ser consistentes con la transmisión de un mensaje de seria preocupación por no terminar violentando injustificadamente, mediante el castigo, la esfera de derechos fundamentales de la persona que eventualmente sea privada de su libertad. Por tanto, debe consistir en una serie de pasos que muestren de parte de las autoridades respectivas un *comportamiento ejemplar*, que muestren la sujeción de aquellas a las reglas derivadas de la modalidad más robusta posible de la noción de «juego limpio» (*fair play*), que muestren un esfuerzo genuino por deslindarse de precipitaciones, abusos, arbitrariedades y de todo tipo de actos que las denigren. «Si termina habiendo sustento para reprochar tu conducta, antes habrás sido tratado como no trataste a tu víctima: con la mayor dosis razonable de consideración y respeto». Esa sería la esencia del mensaje que con el procedimiento debería enviarse a la ciudadanía.

En este punto estamos ya en condiciones de abordar la cuestión del tipo de epistemología jurídica más compatible con la teoría de la justificación del castigo sucintamente presentada. En mi opinión, antes que nada, esta no puede eludir el compromiso de estar al día respecto del estado del arte de las investigaciones sobre la cognición humana, es decir, sobre la forma en que las personas procesamos la información de nuestro entorno a los efectos de lograr determinados objetivos o realizar ciertas tareas.

Un énfasis especial debería hacerse en los estudios acerca de la distinta gama de sesgos de los que podemos ser presa en el procesamiento de la información mencionado anteriormente. Y, en efecto, particular atención merece el denominado sesgo de confirmación, o sea, el sesgo que, con base en los argumentos de este trabajo, he sostenido que puede atribuirse a Laudan al plantear sus diagnósticos general y específico del estado epistémico del sistema acusatorio estadounidense, los cuales, como sabemos, destacan un problema supuestamente muy grave en lo que se refiere a la frecuencia y costos de las absoluciones falsas.

---

derechos de diversa índole, tales como los derechos a la educación, a un trabajo bien remunerado, a una vivienda digna, a la salud, así como los derechos de las personas en prisión. Mediante el reconocimiento y consolidación de los derechos sociales y económicos referidos el Estado estaría contribuyendo a disminuir el riesgo de que las personas recurran a la delincuencia para suplir carencias básicas, para hacer más llevadera una vida marcada por la precariedad y la miseria. Y mediante el reconocimiento y consolidación de los derechos de las personas privadas de su libertad, estaría creando la oportunidad para que estas atraviesen el ciclo antes mencionado que esta teoría concibe al menos como idealmente esperable de un agente moral. En lo que respecta a la labor del legislador, la de confeccionar el derecho penal sustantivo, este debería asegurarse, entre otras cosas, de criminalizar conductas para cuya prohibición claramente existan consideraciones morales, y de que el derecho penal se empleará como última ratio.

[61] En cuya configuración quizá sobra decir que el Estado se juega gran parte de la autoridad moral de la que busca revestirse para reprochar legítimamente a la persona condenada.

Como adelanté en su oportunidad[62], el sesgo en comento consiste en el despliegue a nivel inconsciente de estrategias orientadas a que no se modifique, sino, al contrario, a que solo se robustezca, ante nuestros ojos al menos, la preconcepción o impresión inicial que ya tengamos sobre algún fenómeno o asunto de interés. Entre esas estrategias destacan la de reunir solo información que confirme la preconcepción respectiva (de ahí el nombre de sesgo de confirmación), evitando así el contacto con elementos que puedan refutarla; así como la de sobrevalorar la información confirmatoria recabada, y subvalorar la información que pone en peligro la impresión o hipótesis de partida, que pese a todo haya superado el primer filtro para deshacerse de ella.

Vernos influidos por el sesgo de confirmación es algo que nos ocurre en los más variados contextos ordinarios y profesionales, y de manera mucho más común de lo que quizá sospechemos[63]. No es de extrañar, por tanto, que el jurídico no se escape, incluido el procedimiento penal. De ello dan cuenta, por mencionar algunos de los más significativos, los estudios experimentales de Rassin (2010), Eerland y Rassin (2010) y Ask y otros (2008) que, como manifestaciones del sesgo de confirmación, nos hablan del fenómeno de la ceguera a hipótesis o escenarios alternativos al de la culpabilidad del sospechoso por parte de los sujetos de la muestra (entre ellos, policías y fiscales), o de un escepticismo asimétrico, mismo que es mayor para el caso de información potencialmente exculpatoria que para el caso de información inculpatoria.

---

[62] Véase la sección 4 de este trabajo.

[63] Como apunta Lidén (2018), la psicología cognitiva (pp. 118-121), la psicología de las emociones y de la motivación (pp. 122-129), y la psicología social y organizacional (pp. 129-149), han mostrado, de entrada, que la sola escasez de recursos atencionales aunada a la necesidad de simplificar la abundancia y complejidad de los datos que sería relevante tomar en cuenta para la realización de nuestras actividades, contribuyen a que con regularidad nos concentremos únicamente en ciertas pistas o factores clave, es decir, a que procesemos la información de modo parcial. Con ello, entre otras cosas, logramos que el mundo se ajuste a nuestras expectativas la mayor parte del tiempo, y evitamos disonancias cognitivas que pongan en crisis nuestros presupuestos y la metodología de solución de problemas que hemos desarrollado a lo largo de la vida, formal e informalmente. Cierto es que, con todo y lo selectiva que ya de inicio parece ser la forma en que procesamos la información, esto no siempre conduce a malos resultados. De hecho, el conocimiento que nos hace expertos o especialistas en alguna materia en buena medida descansa en la habilidad de no dejarnos llevar por el «ruido» que normalmente distrae a las personas novicias en el campo de que se trate. Sin embargo, la cara negativa de dicha habilidad se muestra más y más cuando concurren otros factores situacionales y sistémicos que conforman el ambiente social e institucional en el que nos desenvolvemos, tales como la sobrecarga de trabajo, el que tengamos que colaborar con otros, lo cual es frecuente que conduzca a la polarización de puntos de vista y/o a comportamientos complacientes con la corriente a efecto de satisfacer la necesidad de sentirnos pertenecientes a un grupo, la forma en que es evaluado nuestro desempeño, el rol que tengamos en la organización o institución de que se trate, misma que, a su vez, persigue ciertos objetivos y promueve ciertos valores, etcétera. Estos factores provocan en las personas emociones y sentimientos como la angustia, el estrés, la ansiedad, la ira, la frustración, la tristeza o la depresión, los cuales no solo colorean nuestra experiencia, sino que refuerzan ese procesamiento selectivo hasta llevarlo al sesgo descarnado, a comportamientos claramente confirmatorios en su mayoría, con tal de cumplir con las tareas que corresponda, al menos superficialmente.

De modo que no por una cuestión de moda o por mera curiosidad intelectual la epistemología jurídica debería estar al tanto de este aspecto de la cognición humana, sino porque contándose con esa comprensión se adquiere consciencia de que por el solo hecho de cargar al Estado con el peso de tener que probar la culpabilidad de los acusados, el procedimiento penal corre el peligro de degenerarse al punto de transformarse en una maquinaria inquisitoria y punitivista[64], incluso pese a tener un cascarón acusatorio que supuestamente debería complicar el camino para condenar; peligro a cuya contención no puede dejar de contribuir una epistemología que busca que el procedimiento que el Estado implemente le permita revestirse de la autoridad moral para reprochar.

Pero, para que la contribución a la contención referida pueda traducirse en medidas más efectivas de mitigación del sesgo confirmatorio, la epistemología jurídica que propongo tiene también que hacerse una idea acerca de hasta qué punto el sistema de referencia ya ha sido afectado por dicho sesgo[65]. Partir de ese supuesto, o sea, de que

---

[64] En lo que Vigna y Devalpo llaman una *fábrica de culpables.* Véase la sección 1.

[65] En el caso del sistema estadounidense, ese que paradójicamente Laudan describió como uno *demasiado amigable con la absolución*, los estragos del sesgo confirmatorio no parecen ser menores. Baste solo recordar que, de acuerdo con el trabajo de Givelver y Farrell (2012), los jueces coinciden en considerar una cortina de humo, una distracción, una ilusión por la que el jurado no debería dejarse llevar, al hecho de que la defensa del acusado presente una historia alternativa corroborada en algún grado por uno o más testigos, en adición a la declaración en juicio del propio acusado en la que afirme que es inocente, sin que la fiscalía en el contrainterrogatorio presente información alguna sobre los antecedentes penales que aquel pudiera tener (oportunidad que, si los hubiere realmente, no dejaría pasar ningún fiscal una vez que el acusado decide tomar el estrado).

No estoy diciendo, claro, que esos factores probatorios sean indicadores concluyentes de la inocencia del acusado; pero tampoco puede sostenerse plausiblemente que sean detalles ínfimos a los que no hay que poner atención alguna. En todo caso, a efectos de evaluar preliminarmente lo arraigado del sesgo confirmatorio, la atención tiene que ser puesta en el hecho de que los jueces parecen *sistemáticamente* considerar esos factores justo así, como elementos distractores. Si este es el panorama al que nos enfrentamos, la culpabilidad, y no la inocencia, es lo que estaría operando como la presunción de partida (sobre todo respecto de personas negras o hispanas, y/o con antecedentes penales); la cual, además, parece estar siendo concebida como una presunción casi invencible, pues muy poco margen de maniobra le queda a la defensa si en efecto esos factores, en opinión de los jueces, no alcanzan ni siquiera para establecer una duda razonable a favor de su representado.

Ahora bien, dado el desacuerdo de los jueces con una muy alta proporción de las absoluciones pronunciadas por los jurados respectivos, documentado en el análisis de Givelber y Farrell (2012), queda claro que, si hubiese estado en sus manos, habrían condenado a las personas acusadas en esos casos. El problema es que en sus manos está normalmente la cuestión, al ser ellos quienes ratifican cotidianamente los miles y miles de tratos que se dan anualmente entre los fiscales y los acusados, lo cual en Estados Unidos constituye el cauce normal en la actualidad, al punto de que alrededor del 90 por 100 de los casos se resuelve recurriéndose a la figura de la justicia negociada (o *plea bargaining*), y el 95 por 100 de las condenas se obtiene por el mismo medio.

¿Por qué su intervención en la ratificación de los acuerdos es un problema (cuando en principio parece una medida sensata para evitar abusos)? Primero porque, si son acertadas las apreciaciones de Givelber y Farrell (2012), se trata de jueces predispuestos a solo prestar atención a las pruebas de cargo a las que la fiscalía se hubiere referido en el acuerdo respectivo, y en esta fase no hay jurado que les lleve la contraria; segundo porque, además, la persona acusada ha admitido su culpabilidad, lo cual a los ojos de

el sistema de que se trate ya ha sufrido en alguna medida los estragos del sesgo confirmatorio, no es fatalismo, ni apresuramiento, sino una hipótesis realista que toma en cuenta, de un lado, que los procedimientos penales con los que hoy contamos al menos en el mundo occidental se fueron moldeando y consolidando, algunos de ellos durante siglos, en una época en la que nuestros conocimientos sobre el procesamiento a nivel inconsciente de la información eran escasos o, de plano, inexistentes; y, por otro, que la operación del sistema de justicia no está influida solo por las reglas formales u oficiales que le dan forma, sino de modo muy importante también, por los hábitos, convenciones informales, estilos y cultura organizacional que en la práctica se implementan en los precintos policiales, las fiscalías, los laboratorios forenses y los tribunales; lo cual suele pasar desapercibido y, por tanto, desarrollarse por debajo del radar, sobre todo aprovechando los espacios que las reglas formales dejan abiertos para el ejercicio de distintos grados de discreción por parte de los funcionarios respectivos.

Han quedado en el tintero varias cosas por decir sobre el proyecto de epistemología jurídica que aquí he propuesto como alternativa al de Laudan, las cuales dejaré para otras oportunidades[66]. Baste decir, en suma, al menos tres cosas: 1) que se trata de un proyecto esencialmente de *debiasing* o de mitigación del sesgo confirmatorio; 2) que debería tener lazos fuertes de colaboración con las investigaciones en el campo de las condenas erróneas (o *wrongful convictions*)[67], actualmente potenciadas en ese país por la existencia de una base de datos tan importante como el National Registry of Exonerations, y: 3) que dos rubros prioritarios de esta modalidad de la reflexión epistemológico-jurídica, —de hecho, dos pendientes que no pueden postergarse más—, son los mecanismos de justicia negociada[68] y, por supuesto, la investigación del delito[69].

BIBLIOGRAFÍA

Aguilera, E. (2020). Una propuesta de aplicación de la epistemología jurídica en la investigación del delito. En J. Ferrer y C. Vázquez (Coords.), *Del derecho al razonamiento probatorio* (p. 17-44). Marcial Pons.

---

alguien ya inclinado a considerarla culpable, termina por confirmarlo aún más; y tercero porque, como observa Medwed (2012: 52-68), la revisión seria de los acuerdos que en teoría deberían hacer los jueces, dada la sobrecarga de trabajo en la que operan y la fuerte y muy conveniente *presunción de corrección* que hay a favor del proceder de las fiscalías, termina siendo en la práctica una ratificación superficial en la que apenas se cercioran de que la narrativa de los hechos imputados sea más o menos coherente, que no sea difícil encuadrarla en el delito de que se trate, y que las pruebas —solo de cargo recuérdese, y sin que hayan sido practicadas, sino solo enlistadas por el fiscal—, sean mínimamente relevantes.

[66] Algo más sobre este proyecto (y sobre las críticas a la metodología del diagnóstico específico de Laudan y a su propuesta de tratamiento epistemológico), lo desarrollo en (Aguilera, 2025).

[67] Como trabajos sumamente representativos de ese campo, véanse Cook, 2022; Free y Ruesink, 2016; Garret, 2021; Gross, 2022, 2020; Harris, 2012; Medwed, 2017; Morgan, 2023; Norris y otros, 2023; Wojciech y Kremes, 2023.

[68] Véase Lippke, 2011, capítulo 9.

[69] Véanse Aguilera, 2020; Moscatelli, 2023.

Aguilera, E. (2025). *Epistemología jurídica: Cuestiones, debates y propuestas actuales*. Zela.
Ask, K., Rebelius, A., y Granhag, P. (2008). The elasticity of criminal evidence: A moderator of investigator bias. *Journal of Applied Cognitive Psychology*, *22*(9), p. 1245-1259. *https://doi.org/10.1002/acp.1432*
Azaola, E. y Ruíz, M. (2009). *Investigadores de papel: Poder y derechos humanos entre la policía judicial de la Ciudad de México.* Fontamara.
Cassell, P. (2018). Tradeoffs Between Wrongful Convictions and Wrongful Acquittals: Understanding and Avoiding the Risks. *Seton Hall Law Review*, *48*(4), p. 1435-1492.
Cook, K. (2022). *Shattered justice: Crime victims' experiences with wrongful convictions and exonerations*. Rutgers University Press.
Dei Vecchi, D. (2020a). *Los confines pragmáticos del razonamiento probatorio.* Zela.
Dei Vecchi, D. (2020b). Estándares de suficiencia probatoria, moralidad política y costos de error: el núcleo inconsistente de la epistemología jurídica de Larry Laudan. *Doxa, Cuadernos de Filosofía del Derecho*, *43,* pp. 397-426. *https://doi.org/10.14198/DOXA2020.43.15*
Duff, R. A. (1986). *Trials and punishments*. Cambridge University Press.
Duff, R. A. (1996). Penal communications: Recent work on the philosophy of punishment. *Crime and Justice, A review of research*, *20*, pp. 1-96. *https://doi.org/10.1086/449241*
Duff, R. A. (2001). *Punishment, communication, and community*. Oxford University Press.
Eerland, A., y Rassin, E. (2010). Biased evaluation of incriminating and exonerating (non)evidence. *Journal of Psychology, Crime, and Law*, p. 1-8. *https://doi.org/10.1080/1068316X.2010.493889*
Feeney, F., Dill, F., y Weir, A. (1983). *Arrests without conviction. How often they occur and why*. National Institute of Justice.
Findley, K. (2018). Reducing Error in the Criminal Justice System. *Seton Hall Law Review*, *48*(4), p. 1265-1318.
Free, M. y Ruesink, M. (2016). *Wrongful convictions of women: When innocence isn´t enough*. Lynne Rienner Publishers.
Gardiner, G. (2017). In Defence of Reasonable Doubt. *Journal of Applied Philosophy*, *34*(2), pp. 221-241. *https://doi.org/10.1111/japp.12173*
Garret, B. (2021). *Autopsy of a crime lab: Exposing the flaws in forensics.* University of California Press.
Givelber, D. y Farrell, A. (2012). *Not guilty: Are the acquitted innocent?* NYU Press.
Gross, S. (Ed.) (2022). *Race and wrongful convictions in the Unites States.* National Registry of Exonerations.
Hampton, J. (1992). An expressive theory of retribution. En W. Creig (ed.), *Retributivism and its critics* (p. 1-21). V. Steiner Verlag.
Harris, D. (2012). *Failed evidence: Why law enforcement resists science.* New York University Press.
Kalven, H. y Zeisel, H. (1966). *The American Jury.* Little Brown.
Koppi, R. (2018). Comment on Laudan. *Seton Hall Law Review*, *48*(4), p. 1255-1263.
Laudan, L. (2006). *Truth, Error, and the Criminal Law: An Essay in Legal Epistemology*. Cambridge University Press.
Laudan, L. (2013). *Verdad, error y proceso penal: un ensayo sobre epistemología jurídica* (traducción de Carmen Vázquez y Edgar Aguilera). Marcial Pons.
Laudan, L. (2016). *The Law´s Flaws; Rethinking Trial and Errors?* College Publications.
Laudan, L. y Saunders, H. (2009). Re-Thinking the Criminal Standard of Proof: Seeking Consensus about the Utilities of Trial Outcomes. *International Commentary on Evidence*, *7*(2).
Lidén, M. (2018). *Confirmation bias in criminal cases*. [Tesis doctoral, Uppsala University]. *https://www.diva-portal.org/smash/get/diva2:1237959/FULLTEXT01.pdf*
Lidén, M. (2020). Confirmation Bias in Investigations of Core International Crimes: Risk Factors and Quality Control Techniques. En X. Agirre Aranburu, M. Bergsmo, S. De Smet y C. Stahn (eds.), *Quality Control in Criminal Investigation* (p. 461-528). Torkel Opsahl Academic EPublisher.
Lippke, R. (2011). *The ethics of plea bargaining*. Oxford University Press.
Lippke, R. (2016). *Taming the presumption of innocence*. Oxford University Press.
Lippke, R. (2024). *Theorizing legal punishment*. Routledge.

Magaloni, A. (2009). *El ministerio público desde dentro: Rutinas y métodos de trabajo en las agencias del MP*. Centro de Investigación y Docencia Económicas, Documento de Trabajo, 42.
Medwed, D. (2012). *Prosecution complex: America´s race to convict and its impact on the innocent*. New York University Press.
Medwed, D. (2017). *Wrongful convictions and the DNA revolution: Twenty-five years of freeing the innocent.* Cambridge University Press.
Morgan, J. (2023). *Wrongful convictions and forensic science errors: Case studies and root causes*. Taylor and Francis.
Moscatelli, L. (2023). La importancia de la abducción en la etapa de investigación criminal. *Quaestio Facti, Revista Internacional sobre Razonamiento Probatorio*, *5*, p. 125-155.
Norris, R., Hicks, W. y Mullinix, K. (2023). *The politics of innocence: How wrongful convictions shape public opinion.* New York University Press.
Rassin, E. (2010). Blindness to alternative scenarios in evidence evaluation. *Journal of Investigative Psychology and Offender Profiling*, *7*(2), p. 153-163.
Vigna, A. y Devalpo, A. (2010). *Fábrica de culpables: Florence Cassez y otros casos de la injusticia mexicana*. Grijalbo.
Wojciech, J. y Kremes, K. (2023). *Compensation for wrongful convictions: A comparative perspective.* Routledge.
Zalman, M. (2018). «The Anti-Blackstonians». *Seton Hall Law Review*, *48*(4), p. 1319-1433.
Zepeda, G. (2004). *Crimen sin castigo: Procuración de justicia penal y ministerio público en México*. Fondo de Cultura Económica.

Quaestio facti. Revista Internacional sobre Razonamiento Probatorio
Quaestio facti. International Journal on Evidential Legal Reasoning
Sección: Homenaje a Larry Laudan
2026 I 10 pp. 157-185
Madrid, 2025
DOI: 10.33115/udg_bib/qf.i10.23152
Marcial Pons Ediciones Jurídicas y Sociales
© Eunseol Lee
ISSN: 2604-6202
Recibido: 13/05/2025 | Aceptado: 03/11/2025 | Publicado online: 19/12/2025
Editado bajo licencia Reconocimiento 4.0 Internacional de Creative Commons

THE PURSUIT OF TRUTH AND THE JUSTIFICATION IN LEGAL FACT-FINDING: A COMPARATIVE ANALYSIS OF BENTHAM, LAUDAN, AND DAMAŠKA

Eunseol Lee
Université Libre de Bruxelles
eunseol.lee@ulb.be
Université Toulouse Capitole
eunseol.lee@ut-capitole.fr

ABSTRACT: This paper examines the philosophical and legal foundations of the principle of free proof in criminal adjudication through a critical analysis of the theories of Jeremy Bentham, Larry Laudan and Mirjan Damaška. While Bentham and Laudan support the broadest possible inclusion of evidence to promote truth-seeking and reduce judicial error, Damaška emphasizes the institutional, structure and cognitive constraints that limit evidentiary reasoning. Through comparative legal and epistemological inquiry, the paper challenges the assumption that admitting more evidence necessarily advances to discovery of truth. It argues that, although conceptually aligned with the truth-seeking, the free use of evidence is constrained in practice by human biases, procedural safeguards, and epistemic limitations. The study advocates a shift from the uncritical expansion of admissible evidence toward a standard of justified fact-finding where factual conclusions are rationally and legally grounded within a procedurally legitimate framework. Ultimately, the papers calls for a reorientation of evidentiary theory that integrates epistemic justification with the institutional design, ensuring both fairness and the responsible approximation of truth in criminal trials.

KEYWORDS: free evaluation of evidence; admission of evidence; Judicial epistemology; justification of the fact-finding; rules of evidentiary evaluation and judgment rules; exclusionary rules.

SUMMARY: 1. INTRODUCTION.— 2. THE EPISTEMIC FOUNDATIONS AND EVOLUTION OF EVIDENTIARY RULES.— 3. BEYOND RATIONAL OPTIMISM: COMPETING

CONCEPTIONS OF TRUTH IN LEGAL EPISTEMOLOGY: 3.1. Bentham's Reform: Rational Evaluation as a Substitute for Certainty; 3.2. Laudan's error-reduction pragmatism and the problem of justification; 3.3. Damaška's institutional epistemology as a response to both; 3.4. Comparative synthesis: from truth-seeking to justification-seeking; 3.4. Comparative synthesis: from truth-seeking to justification-seeking.— 4. THE DUAL AIMS OF LEGAL EPISTEMOLOGY: FROM TRUTH TO JUSTIFICATION: 4.1. Truth and the Ideal of Free Proof; 4.2. Strict rules of evidence: a non-determining factor in fact-finding.— 5. FURTHER DIRECTIONS: FROM INDIVIDUAL RATIONALITY TO INTERSUBJECTIVE JUSTIFICATION: 5.1. From Objective to Intersubjective Propositional Justification; 5.2. Explanatory Coherence and the Discursive Norm of Assertion; 5.3. Conclusion: Intersubjectivity as an Epistemic and Civic Norm.— 6. CONCLUSION.— REFERENCES

1. INTRODUCTION

The pursuit of truth is widely acknowledged as a central aim of criminal adjudication. Among the mechanisms that serve this objective, the principle of free proof, allowing fact-finders to consider nearly all relevant and lawfully obtained evidence, has gained renewed attention, particularly in inquisitorial systems where judges play a central role in evaluating evidence. In contrast, the principle of regulated proof, which restricts admissibility to protect procedural fairness and prevent bias, reflects a more cautious understanding of rational adjudication. The tension between these two models encapsulates a broader philosophical question: whether truth is better pursued through freedom or through regulation.

This paper revisits the principle of free proof, exclusionary rules, and evidentiary admissibility through the perspectives of Jeremy Bentham, Larry Laudan, and Mirjan Damaška. Bentham, the philosophical founder of free proof, maintained that broad evidentiary inclusion maximizes rational inquiry and aligns adjudication with the discovery of truth. Laudan, a contemporary epistemologist, reformulates this pursuit as an institutional problem of minimizing errors, seeking to design procedural systems that optimally balance wrongful conviction and acquittal. Damaška, in turn, highlights the structural, psychological and epistemological limits that shape how truth is constructed and interpreted within adjudicative institutions.

Despite their shared commitment to truth, a central question remains unresolved: does admitting more evidence necessarily bring us closer to it? Comparative analysis suggests that while the principle of free proof broadens the evidentiary landscape, it does not, by itself, ensure epistemic accuracy or procedural legitimacy. Even in systems where all evidence is freely considered, truth may be reached only by chance or distorted through epistemic and institutional constraints. The problem, therefore, lies not merely in what evidence is admitted, but in how factual conclusions are justified.

Accordingly, this paper argues that the debate on the law of evidence must move beyond the question of admissibility toward a more fundamental inquiry into the

nature of legal proof itself, specifically, what it means for a fact to be legally and epistemically justified. The focus should not rest on whether evidence may be admitted, but on how the fact-finding can be rendered both rationally warranted and procedurally legitimate within the structure of criminal adjudication.

To this end, this study proposes an integrated framework that situates evidentiary reasoning within the broader context of epistemic justification. By placing the works of Bentham, Laudan, and Damaška in dialogue, it argues that the legitimacy of fact-finding depends less on the liberalization of evidentiary rules than on the epistemic form and justification of the facts established. This reorientation, from admissibility to justification, and from individual rationality to intersubjective epistemic reasoning, provides the conceptual foundation for the argument developed in the following chapters.

2. THE EPISTEMIC FOUNDATIONS AND EVOLUTION OF EVIDENTIARY RULES

> The founders (of evidence law) were unaware of limitations on human knowledge, we are told, and believed that certain types of evidence reveal objective truth (Damaška, 2019, p. 27).

According to Damaška (2019), the basic idea behind late medieval rules of evidence rested on two core principles. First, there was a belief that fact-finders could accurately determine objectively occurring facts from the past. Second, there was confidence that certain types of evidence could directly reveal these objective facts. The rules of evidence during this period did not sufficiently consider the rationality or epistemological limitations of fact-finders and assumed that truth could be proven through authoritative rules. Although modern evidence rules face epistemological critiques for obstructing the pursuit of truth, their original intent was to promote truth discovery.

This foundational belief in the possibility of accessing objective truth through certain types of evidence gave rise to a hierarchical structure of evidentiary value. In the medieval era, eyewitness testimony, particularly the testimony of two unimpeachable eyewitnesses (*testes omni exceptione majores*), was regarded as the most reliable form of evidence, treated as definitive rather than probabilistic. Similarly, a defendant's confession, even one obtained under torture, was seen as absolute proof of guilt. Following these were the testimony of a single eyewitness and, lastly, circumstantial evidence, which occupied the weakest position in the hierarchy (Damaška, 2019, pg. 35).

Such hierarchical thinking did not emerge in isolation but rather evolved from longstanding legal traditions (pg. 35-36). Notably, a similar structure can be found in Roman law, where direct evidence was prioritized over circumstantial evidence. The logic was straightforward: while circumstantial evidence required interpretation, direct evidence appealed to the senses and was presumed to mirror reality more

faithfully (p.18; p. 36).[1] For instance, while the presence of a knife in a victim's home might suggest but not prove guilt, a credible eyewitness's account of the crime was considered to offer stronger inferential certainty. These categorizations aimed to reduce the margin of error in the judicial process.

Closely aligned with this evidentiary hierarchy were the methods of trial by ordeal and trial by combat, which were also understood as mechanisms for uncovering truth.[2] These procedures were justified through theological reasoning that presumed divine intervention: God would grant victory to the righteous and expose falsehoods through suffering (Haack, 2014, p. 2). Thus, the judicial process was not only evidentiary but also cosmological, rooted in metaphysical assumptions about divine justice. These practices reveal how medieval legal epistemology was embedded in a worldview that denied the very possibility of uncertainty. Truth was conceived as singular, absolute, and divinely accessible. Accordingly, medieval evidence law did not aim at an approximate or probabilistic understanding of facts but sought to identify the precise sequence of past events through authoritative or divinely sanctioned means.

As legal systems entered the modern era, however, this worldview began to change. Courts and scholars gradually acknowledged the limits of procedural mechanisms and of human epistemic capacity. This acknowledgment did not create a new concept of truth but transformed its meaning: truth came to be understood not as divine certainty but as the most accurate reconstruction of past events that is humanly and institutionally attainable. This paradigmatic shift was not merely theoretical; it had direct implications for the structure of evidence law. As rationalism and empirical thought gained traction, so too did skepticism toward irrational forms of evidence.[3] Practices such as confession under duress, divine judgment, or deference to the social status of a witness lost legal and moral credibility. In their place emerged a growing emphasis on logical coherence, empirical verification, and procedural transparency,

---

[1] Damaška, M. (2019). *Evaluation of Evidence: Premodern and Modern Approaches*. Cambridge University Press p. 18, p. 36. Katherine Tachau cites Ockham's razor to point out that perception shows objects more directly than inference. "Yet the very point of his [Ockham's] insistence that all adjudicative and abstractive cognitions presuppose intuitions is that the latter allow us to know—immediately and with certitude—that an object exists, rather than merely to infer it". See Tachau K. (1988).

[2] See Neilson, G. (1890). *Trial by Combat,* Williams and Norgate. According to Susan Haack, "to pick up a ring from the bottom of a cauldron of boiling water, and his arm would later be checked to determine whether it had healed cleanly or had festered—which supposedly showed that he was guilty; in trial by combat, the two parties to a case would literally fight it out".

[3] For example, the practice of evaluating the value of testimony differently based on the social status or gender of the witness does not meet reasonable standards. It is no longer reasonable to consider the testimony of male witnesses as stronger evidence because women are deemed less capable than men in providing eyewitness testimony. The argument that male testimony is more credible than female testimony can be found below. Damaška (2019, p.61) adds examples to support this argument, considering several conflicting testimonies. He cites cases where the testimony of a woman who is a respectable property owner but lives a wicked life conflicts with that of a man of low status but virtuous character, and cases where the testimony of three men conflicts with that of four women. In fact, in Saudi Arabia, male testimony is considered twice as important as female testimony. See Haack (2014, p. 3).

elements foundational to modern rules of admissibility. The intellectual underpinnings of this shift were heavily influenced by modern philosophy, particularly Enlightenment rationalism and empiricism (Damaška, 1997, p. 20). Thinkers such as Jeremy Bentham played a critical role in challenging the validity of traditional evidentiary hierarchies. Bentham condemned blanket rules that assigned predetermined weight to specific types of evidence, arguing that such practices were arbitrary and obstructed the rational understanding of individual cases (Damaška, 2019, p.1). As belief in divine truth waned, faith in human reason as the primary instrument of legal truth-finding took root.

This evolution also led to the recognition that modern legal systems must pursue truth within the constraints of fairness, institutional capacity, and human rights. Evidence law began to incorporate non-epistemic goals alongside the pursuit of truth: ensuring due process, safeguarding dignity, promoting procedural efficiency, and maintaining public trust in the legal system. Rather than conflicting with the goal of truth discovery, these secondary aims were increasingly seen as reinforcing its legitimacy and reliability.

From the medieval reliance on divine revelation to the Enlightenment's faith in rational evaluation, the history of evidence law traces humanity's shifting conception of truth. As procedural formalism gave way to the ideal of free proof, reason replaced revelation as the foundation of legal epistemology. Yet this transformation did not abolish uncertainty, it merely relocated it. Modern law now recognizes that truth-seeking is inseparable from fairness, institutional design, and cognitive limitation. The challenge for contemporary evidence theory is therefore not simply to expand evidentiary freedom, but to determine how factual conclusions can be responsibly justified within these human and procedural constraints.

3. BEYOND RATIONAL OPTIMISM: COMPETING CONCEPTIONS OF TRUTH IN LEGAL EPISTEMOLOGY

The history of evidence law reflects a persistent tension between the aspiration to discover truth and the acknowledgment of human and institutional limits. Against this backdrop, the following sections examine three distinct responses to the problem of truth in adjudication. Jeremy Bentham represents the classical rationalist faith in evidentiary freedom, the conviction that reason, if unshackled from formal constraint, will naturally converge upon truth. Larry Laudan reformulates this ideal through a pragmatic and probabilistic lens, transforming Bentham's rationalism into an empirically grounded project of error reduction. Mirjan Damaška, in contrast, challenges both by situating legal truth within the institutional conditions of justified belief. Together, these thinkers trace the movement of legal epistemology from the optimism of pure rationality to the realism of epistemic justification.

This chapter therefore explores not only competing conceptions of truth but also competing visions of legitimacy in adjudication. For Bentham, legitimacy derives from the rectitude of decision, the factual accuracy of verdicts. For Laudan, it stems from procedural efficiency in minimizing error. For Damaška, it lies in the rational justifiability of belief within institutional and moral constraints. By comparing these frameworks, the chapter seeks to illuminate the broader transformation of legal epistemology: from a system devoted to truth as certainty to one grounded in justification as responsibility.

3.1. Bentham's Reform: Rational Evaluation as a Substitute for Certainty

The modern transformation of evidentiary thought began with a decisive break from the hierarchical formalism that had governed medieval and early modern law. Jeremy Bentham (1843) denounced this system as irrational for its reliance on rigid categories of proof and predetermined evidentiary weights. In his utilitarian reform, Bentham replaced divine certainty with the authority of human reason. All relevant evidence, he argued, should be admissible and freely assessed by the fact-finder, whose rational deliberation would serve as the principal mechanism for truth discovery. For Bentham (, as cited in Twining, 2019, p. 27), justice depended not on the social status of witnesses or the ritual of procedure but on what he termed the *rectitude of decision,* the factual accuracy of judicial outcomes. Excluding probative evidence, in Bentham's view (Bentham, 1847, p. 384), was tantamount to excluding justice itself. A decision is just only insofar as it corresponds to substantive truth; any rule that obstructs access to relevant facts therefore undermines the very objective of law. Procedural fairness that impedes truth-seeking represents, for him, a corruption of utility (Haak, 2014, p.5). He famously likened such fairness to "the foxhunter's reason", (Bentham, 1827) arguing that granting procedural advantages to defendants through rigid exclusionary rules was akin to giving a fox a head start from hunters. If the ultimate purpose of adjudication was the discovery of truth, then shielding relevant evidence under the guise of fairness merely served "Injustice and her handmaid Falsehood" (Bentham, 1827).

According to Bentham, "the direct end of adjective law is rectitude of decision, that is, the correct application of valid laws (presumed to be consonant with utility) to true facts" (Bentham, as cited in Twining, 2006, p. 41). While a distinction may be drawn between rectitude of decision and the delivery of just outcomes, Bentham (Bentham, as cited in Twining, 1985, p. 28) regarded them as inseparable: justice required that facts be accurately established and that legal consequences applied to them be fair and proportionate. Rigid evidentiary rules, he argued, were counterproductive precisely because they excluded probative material *a priori* without meaningful examination. This subverted the trial's purpose and limited the fact-finder's capacity to determine the truth. As he observed, any system that excludes potentially relevant

evidence merely because it belongs to a "category of imperfect evidence" is irrational and detrimental to justice (Schauer, 2009). All relevant material, he insisted, should be presumptively admissible, and exclusion justified only to prevent greater injustice, such as excessive delay or expense (Schauer, 2009, p. 22). Categorical exclusionary rules and precedent, which he called the "path of constant error", (Bentham, 1843) replaced reasoned utility with mechanical ritual. The law, he maintained, must remain empirical and utility-oriented, not formalistic or precedent-bound.

Bentham's theory of evidence thus marked a decisive break from the hierarchical structures of proof that had characterized premodern jurisprudence. Rejecting predetermined evidentiary weights, he maintained that all relevant material should be admissible and freely assessed by a rational fact-finder. Within his utilitarian framework, justice could be achieved only when factual conclusions corresponded to reality, for factual correctness constituted the moral foundation of legal utility. His reform aimed to transform adjudication from a ritualized formality into a rational and empirical inquiry: where medieval epistemology sought divine revelation, Bentham sought verification through human reason.

His faith in unrestrained rationality reflected the broader epistemic optimism of the Enlightenment (De Champs, 2015). Rationalism and empiricism together cultivated the belief that disciplined observation and logical analysis could reconstruct reality with precision (Twining, 1985, pg. 3).[4] The fact-finder, modeled on the Enlightenment scientist, was imagined as an objective observer capable of processing information without distortion. This intellectual climate replaced divine infallibility with the supposed infallibility of reason. The Enlightenment slogan *sapere aude*, "dare to know", embodied a new confidence in human rational agency. Enlightenment epistemology thus replaced the theology of revelation with what might be called the theology of reason, exchanging one form of certainty for another. This rationalist heritage reached its fullest legal expression in the modern principle of free proof, particularly influential in continental systems, which empowers judges to admit and evaluate all relevant evidence without formal restriction. It rests on two assumptions: that factual reality is fully accessible through evidence, and that unrestrained rational evaluation can reliably uncover it. This model clearly bears Bentham's imprint: adjudication reimagined as a reason-based process unburdened by ritual hierarchy or procedural formality.

Yet this rationalist optimism also carried a critical blind spot. The doctrine of free proof overlooks the epistemic, moral, and institutional limits that shape human judgment, assuming that unregulated reasoning is inherently self-correcting when in fact it often amplifies bias and uncertainty. Modern evidentiary law, particularly within adversarial systems, has since tempered this optimism by embedding procedural safeguards, rules of admissibility, exclusionary doctrines, and standards of proof. These mechanisms are not merely moral constraints but epistemic structures

[4] Bentham's system of procedure are based on utility, empiricism, and common-sense reasoning.

designed to preserve the integrity of judicial reasoning. Excluding coerced confessions, protecting the right to silence, and suppressing unlawfully obtained evidence serve not only fairness but truth itself, ensuring that conclusions remain rationally defensible within human limits.

At the core of Bentham's framework lies the principle of rational inclusion, (Bentham 1827) the belief that truth emerges from the comprehensive evaluation of evidence. Yet behavioral decision research has demonstrated that human reasoning is vulnerable to bias, overconfidence, and cognitive overload.[5] Unrestricted evidence can therefore diminish, rather than enhance, accuracy. Bentham's reform thus replaced one absolutism with another, substituting formal hierarchy with faith in reason while neglecting the need for epistemic justification: a structured account of how conclusions can be warranted within institutional and cognitive limits. His model promised liberation from formalism but failed to recognize the realities of adjudication, where rational evaluation alone cannot guarantee truth.

Bentham's utilitarian rationalism therefore marks both the culmination and the limit of Enlightenment epistemic optimism.[6] While it freed evidentiary reasoning from dogmatic hierarchy, it overlooked the fragility of human judgment. A contemporary philosophy of evidence must thus move beyond this rationalist legacy, seeking not unbounded freedom of proof but a system of justified fact-finding, one that aligns truth-seeking with epistemic responsibility while preserving the moral and institutional integrity of law.

3.2. Laudan's error-reduction pragmatism and the problem of justification

Larry Laudan represents one of the most sophisticated modern defenders of the epistemic ideal underlying the principle of free proof. In *Truth, Error, and Criminal Law* (Laudan, 2006), he reinterprets Bentham's rationalism through an empirically grounded model of error reduction. For Laudan, the criminal trial's primary purpose is not the ceremonial enforcement of procedural fairness but the epistemic task of distinguishing the guilty from the innocent. A legal system that tolerates wrongful convictions or acquittals without actively seeking to minimize them, he argues, cannot command moral or epistemic legitimacy. Accuracy in adjudication, rather than procedural observance, determines whether a verdict deserves public respect (2006, p. 1-2).

In this framework, the criminal trial becomes an epistemic engine designed to optimize the ratio between true and false outcomes (Laudan, 2006). The value of any procedural safeguard, evidentiary rule, or exclusionary doctrine must therefore be assessed by its contribution to truth discovery (Laudan, 2006, p. 29). This stance

[5] See Kahneman. (2011).

[6] We can even find a utilitarian rational in *On Crimes and Punishments of Beccaria*. See, Beccaria. (1995). §4, p. 16).

echoes the dictum of the U.S. Supreme Court that "the fundamental purpose of the trial is the search for truth" (Tehan vs U.S. 1966). Rules that do not improve accuracy should, in his view, be revised or discarded. Procedural fairness, for Laudan, is not an independent good but an instrumental one, justified only insofar as it promotes epistemic accuracy.

He is particularly critical of rules that exclude highly probative evidence in the name of fairness. Like Bentham, Laudan views such exclusions as epistemic obstacles disguised as moral safeguards. The right to silence, for example, prevents fact-finders from considering relevant behavioral evidence, thereby limiting their capacity for rational inference (Laudan, 2006, p.117). Before Griffin v. California (1965), juries were free to draw inferences from a defendant's refusal to testify (Laudan, 2006, p.170). The Court's ruling that such inferences violate the Fifth Amendment, Laudan argues, compels jurors to suppress reasoning that would otherwise occur naturally. The prohibition thus represents an epistemic loss rather than a moral victory: it enforces blindness where epistemic evaluation should take place. By shielding defendants from the epistemic consequences of their own choices, the system sacrifices rational transparency for symbolic fairness.

This criticism extends to exclusionary rules more generally. Laudan questions whether doctrines that bar relevant evidence, such as certain forms of hearsay, truly serve justice when they obstruct the fact-finder's access to truth. Under the guise of protecting rights, these rules often produce the perverse effect of protecting the guilty as well as the innocent. The result, in his view, is less truth and more error. Bentham had made a similar point two centuries earlier: when the objectives of fairness and accuracy conflict, the principle of utility should decide in favor of rectitude of decision. Laudan updates this utilitarian logic for modern epistemology, arguing that procedural safeguards must be weighed against their epistemic costs. If a rule increases the risk of false acquittals or convictions, it cannot claim unconditional moral authority.

To adjudicate such trade-offs, Laudan develops what he calls a meta-epistemology of criminal law: a higher-order framework of meta-rules that determine whether legal procedures contribute to or detract from epistemic goals.[7] A meta-rule, as he defines it, provides a criterion for assessing the truth-conduciveness of procedural norms. The fundamental concept of meta-rules is based on the following idea:

> Having settled on the appropriate level of bias, however, we should let the rest of the system function as an epistemically respectable engine, that is, as a viable, error-reducing, distributionally neutral tool for investigation that tries unstintingly to find the truth (Laudan, 2006, p. 144).

Rather than taking rules as axiomatic, Laudan urges that they be tested empirically: do they increase or decrease the probability that verdicts correspond to reality?

[7] Laudan. (2006). "(…) any given procedure or evidence-admitting or evidence-excluding practice does, in fact, further epistemic ends or whether it thwart them".

In this sense, the meta-epistemology of criminal law replaces dogmatic adherence to legal tradition with an empirically informed form of epistemic governance. Procedures are legitimate only insofar as they demonstrably reduce error while maintaining minimal moral constraints.

Central to this framework is Laudan's proposal for quantifying the relationship between epistemic and social values through his Standard of Proof (SoP) formula: SoP = 1 / (1 + 1/n) (Laudan, 2006, p. 72). Here, *n* represents the relative moral severity that society attributes to wrongful conviction compared with wrongful acquittal. The higher the moral cost of convicting the innocent, the greater the evidentiary threshold required to convict. If society regards wrongful conviction as ten times more serious than wrongful acquittal (*n = 10*), the standard of proof adjusts accordingly. Through this model, Laudan operationalizes moral aversion into procedural calibration. The appropriate level of bias is thereby determined by social consensus rather than metaphysical certainty. Once that bias is fixed, he insists, "the rest of the system should function as an epistemically respectable engine, that is, as a viable, error-reducing, distributionally neutral tool for investigation that tries unstintingly to find the truth" (2006, p. 72).

This model introduces a form of institutional realism absent in Bentham's thought. Whereas Bentham presupposed ideal reasoners capable of objectively processing unlimited evidence, Laudan grounds epistemic ambition within epistemic, moral, and institutional limits. He recognizes that truth-seeking must be balanced against the capacities and constraints of real institutions: the bounded rationality of jurors, the asymmetry of appellate review, and the moral costs of judicial error. For instance, Laudan emphasizes that most legal systems allow appeals from convictions but not from acquittals, creating a structural bias toward false acquittals over time (2006, p. 14). This asymmetry, he argues, gradually erodes the system's capacity to discover truth, as errors of acquittal accumulate unchecked.[8] The legitimacy of criminal adjudication therefore depends on a continual re-evaluation of its institutional biases (2006, p.141).

Despite these advances, Laudan's framework remains incomplete at the level of epistemic justification. By equating truth with error minimization, he transforms epistemology into a managerial science of uncertainty. His meta-rules explain when procedures are efficient but not why the beliefs they generate are justified. Truth, though necessary, is not sufficient for epistemic justification. A verdict may be accurate by coincidence yet still lack rational warrant.

This limitation reveals a deeper continuity with Bentham. Both conceive rationality instrumentally rather than normatively: as a means of producing correct

[8] Although Laudan did not mention it, the tendency to hand down not guilty verdicts more easily than guilty verdicts in court is also related to the economics of litigation. When a defendant is found not guilty, the overall litigation costs for a case are lower than when a defendant is found guilty, as the defendant cannot appeal.

outcomes, not as a standard for justified belief. Bentham sought truth through unrestricted inclusion; Laudan seeks it through calibrated optimization. In both, reason functions as an epistemic mechanism rather than a justificatory practice. The result is what may be termed epistemic instrumentalism, the view that truth matters only as procedural efficiency, detached from the rational warrant of belief. Recognizing this limitation reorients legal epistemology from procedural efficiency toward epistemic responsibility. The legitimacy of adjudication cannot rest solely on reducing error; it must also rest on the rational warrant of its conclusions. A judicial system that merely manages uncertainty does not necessarily produce justified beliefs about the past. The central question, therefore, is not how to make verdicts more accurate, but how to ensure they are epistemically justified.

Laudan's contribution, in this sense, serves as a bridge rather than a destination. His empirical realism provides the groundwork for an epistemically responsible theory of adjudication, but his framework remains procedural and instrumental. The next step must move beyond the management of error toward a conception of institutional epistemology that treats justification, not mere accuracy, as the foundation of legitimacy. It is precisely this transition that Damaška's model, and the argument developed in the next section, seek to accomplish.

3.3. Damaška's institutional epistemology as a response to both

Mirjan Damaška offers a fundamentally different vision of legal epistemology one that departs equally from Bentham's rationalist optimism and Laudan's procedural instrumentalism. While both thinkers locate truth in the rational or procedural capacities of fact-finders, Damaška shifts the focus to the epistemic architecture of adjudication itself. Legal truth, in his account, is not the product of unbounded reasoning or statistical calibration but a form of institutional justification: a process through which fact-finding acquires justification within epistemic, procedural, and moral limits.

Unlike Bentham, who trusted in unrestricted access to evidence, or Laudan, who sought to minimize error through procedural optimization, Damaška treats evidentiary rules as epistemic instruments.[9] They structure and discipline reasoning rather than impede it. The central question is not how much evidence can be admitted but under what conditions belief-formation can remain rational and justified. Fact-finders are bounded agents whose reasoning must be shaped and constrained by institutional design.

In *Evidence Law Adrift* (Damaška, 1997) and related works, Damaška reconceptualizes adjudication as a form of institutional epistemic decision-making. Truth in law depends not only on individual rationality but on the system that organizes

[9] See Damaška, 2019, (ch. 2).

and validates that rationality (p. 5). Procedural rules, standards of proof, burdens of persuasion, and exclusionary doctrines perform epistemic functions: they preserve the integrity of reasoning against distortion and bias. The law of evidence constitutes a system of epistemic calibration, a framework for structuring judgment so that belief-formation within the institution remains both rational and legitimate.[10] This reconceptualization reframes adjudication as a process of justified belief-formation rather than absolute truth-seeking. Since no epistemic system can guarantee truth, the trial's legitimacy depends on whether its conclusions are rationally and institutionally justified (Damaška, 1997, p. 121-123). Truth becomes a regulative ideal, guiding inquiry while acknowledging fallibility. The task of adjudication is therefore to produce beliefs that are not only accurate but epistemically defensible.

Damaška's view highlights the dual nature of the law of evidence: it is both epistemic and normative. Its aim is to ensure that conclusions are justified under human and institutional constraints. Excluding coerced confessions or unlawfully obtained evidence, for example, is not merely a moral act but an epistemic one: such evidence undermines the very conditions of trustworthy belief-formation.[11] Procedural safeguards thus embody epistemic humility, the recognition that cognition must be disciplined by structure to remain rational.

Critics of exclusionary rules argue that they obstruct truth by limiting available information. From a Benthamite perspective, more evidence means more truth. Damaška, however, contends that this view oversimplifies the epistemic landscape. The problem is not whether evidence is admitted or excluded but how all evidence, lawful or unlawful, inevitably shapes the reasoning process. Even when judges exclu-

[10] "But the internal fission of the tribunal explains more than the need for formal instructions in advance of jury deliberation. Under the loupe of comparative analysis it soon becomes clear that this fission also explains several striking features of common law evidence that have remained unidentified so far. This is most clearly the case with much of the law of presumptions, as well as with several doctrines governing the burden and the standards of proof. In all Anglo-American jurisdictions, these subjects are deeply entangled in normative webs. But where common law goes into Lilliputian detail, Continental law employs mostly broad regulative principles — especially in the sphere of criminal justice. To a degree, this contrast can be attributed to the already mentioned disparities in ordering the law that prevail in the two branches of the Western legal tradition. And, as I shall argue in Chapter 4, the contrast is also related to different patterns of allocating control over procedural action. A factor of immediate interest, however, is the unequal difficulty faced by unitary and bifurcated courts in attuning amateur adjudicators to prevailing fact-finding methods and conventions" (Damaška, 1997, p. 53).

[11] "More peculiar and also more important are common law rules rejecting probative material on the theory that it might be overestimated or on the theory that its probative value is overshadowed by its "prejudicial" capacity — its capacity to unfairly predispose the trier of fact toward a particular outcome. The hearsay rule is by far the best known example of exclusion on the first of these two grounds. Although countries outside of the common law's compass are not unaware of hearsay dangers, their reaction to them seldom assumes the form of exclusionary rules.22 Where it does, as is sporadically the case in criminal procedure, the embrace of the exclusionary option is rooted as much in due process values as it is in the desire to protect the adjudicator from unreliable information" (Damaška, 197, p. 15).

de inadmissible material, its cognitive imprint may persist.[12] Hence, the real challenge lies not in regulating access to evidence but in understanding how reasoning itself is influenced and justified within procedural form.

To illustrate, Damaška (2022, p.199) describes the dilemma of the "hypothetical judge". Even when a judge correctly excludes illegally obtained but highly probative evidence, prior exposure to it may unconsciously affect evaluation of the remaining record. The decision must be justified as if the tainted evidence were unseen, yet the mind cannot fully unsee it. This example exposes a fundamental epistemic tension: compliance with form cannot ensure epistemic purity. Human judgment is inherently holistic and context-sensitive, vulnerable to bias and contamination despite rational discipline.

This insight undermines both Bentham's and Laudan's assumptions that either rational inclusion or procedural optimization can secure truth. Neither free nor constrained proof escapes the fact that legal truth is the product of bounded human reasoning. The search for truth must therefore be reconceived as the justification of belief within cognitive and procedural limits. Rules of evidence are not external constraints but integral components of this justificatory structure. When a court excludes probative yet tainted evidence, it enacts epistemic self-restraint, an acknowledgment that some information, however revealing, cannot ground justified belief without compromising moral or cognitive integrity.

The core issue, then, is not whether free proof or regulated proof is superior, but how reasoning can remain epistemically accountable within institutional boundaries. For Damaška (2022), the legitimacy of verdicts depends less on their correspondence to reality than on their capacity for rational and public justification. Legal truth is always partial and constructed: an approximation achieved through disciplined reasoning under constraint. This shift from truth-seeking to justification-seeking marks Damaška's major contribution to legal epistemology. Against Bentham, he rejects the fiction of the omniscient fact-finder; against Laudan, he rejects the reduction of epistemology to probabilistic management of error. Instead, he envisions adjudication as a moral-epistemic practice of responsible belief-formation. The legitimacy of law thus depends not on the unbounded pursuit of truth or the efficient reduction of error but on the cultivation of epistemically justified belief within structured institutions.

Damaška's institutional epistemology provides the conceptual foundation for the broader argument of this study: that the legitimacy of adjudication rests on epistemic justification, not evidentiary freedom. By embedding reasoning within procedural form, he reveals that epistemic integrity arises not from the quantity of evidence admitted but from the quality of reasoning permitted. Legal truth, in this light, is

[12] Research related to the influence of illegal evidence can be found in the following source (Damaška, 1997, p.15). See also Stably, *et al.* (2006) p. 469-492.

not an unfiltered mirror of reality but a disciplined construction, an aspirational yet justified belief that honors both fairness and the limits of human cognition. Only when legal systems internalize these limits can they achieve genuine epistemic responsibility.

3.4. Comparative synthesis: from truth-seeking to justification-seeking

The preceding discussion traces an intellectual trajectory in the epistemology of legal proof, from Bentham's rationalist optimism, through Laudan's pragmatic instrumentalism, to Damaška's institutional realism. Each thinker addresses a distinct epistemic challenge: Bentham the arbitrariness of formal proof, Laudan the inefficiency of procedural moralism, and Damaška the illusion that reason or procedure alone can guarantee truth. Collectively, they mark the transition from truth-seeking as the central criterion of evidentiary legitimacy to a more complex ideal of epistemic justification within institutional limits.

Bentham's project begins with liberation. Rejecting the hierarchical formalism of medieval law, he envisioned adjudication as a rational and empirical process freed from ritual and authority. By allowing all relevant evidence to be heard, he sought to democratize fact-finding: truth, he believed, would emerge from the open contest of facts. Yet this rationalism rested on an idealized fact-finder, one immune to bias[13] and cognitive limitation.[14] In reality, unbounded evidence can overwhelm reasoning and reinforce preexisting beliefs. Modern insight reveals that Bentham's faith in rational inclusion risks substituting one form of absolutism for another: the infallibility of reason for the infallibility of divine order.

Laudan inherits Bentham's commitment to truth but transforms it through empirical realism. For him, the trial's purpose is to minimize epistemic error rather than to secure absolute truth. His meta-epistemological framework measures the value of rules and procedures by their capacity to reduce wrongful convictions and acquittals. In this model, procedural fairness is instrumental, justified only insofar as it contributes to epistemic accuracy. By formalizing trade-offs between moral and epistemic costs, Laudan introduces a pragmatic system of calibrated truth-seeking. Yet this very calibration transforms truth into a probabilistic ideal, something managed statistically rather than justified rationally. The legitimacy of verdicts becomes a matter of efficiency rather than epistemic warrant.

Damaška transcends this tension by shifting attention from outcomes to the structure of justification itself. Neither unbounded reasoning nor optimized procedure, he argues, can yield justified belief unless institutional conditions make re-

[13] Guthrie., Rachlinski and Wistrich. (2001, p. 777-830).
[14] See Busey & Loftus (2007, p. 111-117).

asoning itself accountable. His concept of institutional epistemology redefines adjudication as a structured practice of belief-formation that accepts human fallibility while imposing procedural safeguards to sustain rational defensibility. Evidentiary rules, on this account, do not merely regulate access to information; they stabilize reasoning, ensuring that verdicts remain justified both epistemically and normatively. Truth here functions as a regulative ideal, an aspiration pursued through disciplined, public reasoning rather than an endpoint to be reached.

Seen together, Bentham privileges epistemic access, Laudan epistemic optimization, and Damaška epistemic justification. This movement, from inclusion to calibration to justification, mirrors a broader philosophical shift from Enlightenment rationalism to modern epistemic fallibilism. Where rationalism sought certainty through method, fallibilism seeks legitimacy through justification. Adjudication, accordingly, must integrate truth and fairness within a framework of epistemic responsibility: it must produce not only accurate outcomes but beliefs that can be rationally defended within epistemic and institutional limits.

In this light, the pursuit of truth becomes an ethical as well as an epistemic commitment. The law's integrity lies not in achieving perfect correspondence with reality but in reasoning well within its own constraints. The historical movement from Bentham to Laudan to Damaška thus signifies more than a change in evidentiary doctrine, it marks a transformation in the very conception of legal truth. The ideal of absolute discovery yields to the discipline of justified approximation, and epistemic integrity replaces epistemic optimism as the defining virtue of modern adjudication.

4. THE DUAL AIMS OF LEGAL EPISTEMOLOGY: FROM TRUTH TO JUSTIFICATION

The doctrine of free proof emerged as an Enlightenment project to liberate legal reasoning from dogmatic formalism and to empower the fact-finder's rational autonomy. It promised that the freer the use of evidence, the closer justice would come to truth. Bentham and, later, Laudan, each in their own way, sought to realize this aspiration, Bentham through utilitarian rational inclusion, Laudan through an empirically grounded epistemology of error reduction. Both regarded the legitimacy of adjudication as dependent on the accuracy of its outcomes.

This chapter builds on, rather than rejects, that project. Laudan's insistence that legal epistemology be assessed in truth-conducive rather than moralistic terms was a decisive corrective to earlier proceduralism. Yet epistemic justification cannot be exhausted by accuracy or error management. It also depends on how factual beliefs are formed and justified within the epistemic and institutional limits of adjudication. The shift, therefore, is not away from truth but toward a richer conception of justification, one that measures legitimacy by the quality of reasoning rather than by outcomes alone.

The central claim advanced here is that the legitimacy of adjudication rests on epistemic justification: the disciplined process by which factual conclusions are warranted within human and institutional constraints. Truth remains the ultimate aspiration, but justice demands more, it requires that the pursuit of truth proceed under the conditions of reasoned justification. The following sections develop this argument by reframing evidentiary freedom, cognitive limitation, and procedural fairness as instruments of justified belief rather than as obstacles to truth.

4.1. Truth and the Ideal of Free Proof

Proponents of the free use of evidence ultimately seek to discover the truth about the past events of a trial. While "truth" is theorized in different ways, it generally denotes a comprehensive and accurate understanding of the relevant facts, including not only whether a criminal act occurred but also why the defendant acted, the circumstances of decision-making, and any preparatory conduct. Truth, so conceived, cannot be reduced to identifying the "true culprit"; it requires a nuanced, context-sensitive reconstruction. At the same time, the fact-finder, positioned as a third party reconstructing past events from the parties' submission, can never attain absolute certainty. Accurate fact determination (the "discovery of truth" in Bentham's and Laudan's sense of probabilistic terms (Laudan, 2006, p. 20-21)) is an epistemically demanding enterprise irrespective of the governing rules of evidence. For that reason, the pursuit of truth in criminal adjudication is best understood as an ultimate, though often unattainable, ideal.

Whether sweeping away admissibility constraints *increases* truth discovery remains contested. A common claim is that rational deliberation fares best when the tribunal sees the whole picture. Yet comparative practice gives no clear confirmation. Systems influenced by free-proof traditions (e.g., France or South Korea) empower judges to admit and weigh nearly any relevant item; U.S. courts, by contrast, deploy dense admissibility doctrines tied to lay-jury fact-finding. Differences in appellate asymmetries (routine appeals from convictions but rarely from acquittals), investigative powers, standards of proof, and decision-maker type (jury versus professional judge) make reliable cross-system comparisons elusive. In this methodological fog, the syllogism "more admissible evidence → more truth" is unproven. At most, free proof is not a necessary condition for discovering truth.

A familiar illustration refines this point. In a slightly modified version of Michael Pardo's example (2005, p.322), two police officers secretly place cocaine in a car and later perjure themselves at trial, claiming they found it with the driver's consent. The defendant, fearing sentence enhancement if his record is exposed, declines to testify. The fact-finder convicts based on the officers' testimony. Ironically, the defendant did in fact possess cocaine, but only he knew this. The trial disclosed a distorted causal narrative built on false testimony that happened to coincide with the true conclusion. Although the verdict ("the defendant possessed cocaine") was accurate,

its accuracy was accidental and thus epistemically indistinguishable from error. The belief was true but unjustified; it rested on unreliable inferences from tainted testimony rather than on warranted reasoning from admissible, trustworthy evidence.

If the defendant had simply testified truthfully, the epistemic problem in Pardo's example would not arise, suggesting that Laudan's epistemological program remains intact. However, this misses the central claim of the present argument: the issue is not whether Laudan's framework is mistaken, but whether his conception of epistemic rationality should remain confined to the cognitive domain of individual reasoning. The proposal here is to extend Laudan's epistemic project toward intersubjective justification, a model in which the justification of factual conclusions is not only internally rational but also externally shareable, publicly assessable, and communicatively legitimate within the legal forum.

It is therefore essential to evaluate not only whether a conclusion is accurate, but also the legitimacy of the established facts, the validity and manner of using evidence, and the role of chance. In scenarios like Pardo's, a conscientious fact-finder may reasonably credit police testimony and draw an inference of possession; the defendant may exercise the right to remain silent for independent reasons[15]; and the physical presence of cocaine in the car is, on its face, strongly probative. A conviction might appear "justified" in a narrow, record-bound sense. Yet, because the route to that conclusion was epistemically defective, resting on perjury and contaminated inference, the result lacks justificatory legitimacy even if it is descriptively correct. Accuracy without warrant is not an epistemically acceptable endpoint for adjudication.

This reframing has two immediate consequences. First, the central inquiry shifts from *what* enters the record to *how* reasons derived from that record can be made publicly defensible against challenge. Second, exclusions and reliability screens are no longer cast as moralistic obstacles, but as instruments that stabilize reasoning under cognitive and institutional constraints. Even in open regimes, fact-finders see only party-selected slices of reality; they confront asymmetric resources and strategic curation. Human judgment is further bounded by bias, overconfidence, and cognitive load. Without guardrails, more evidence can cause noise, amplify anchors, or create spurious coherence.

---

[15] Salinas v. Texas (2013). This case concerns whether silence during a police interview is protected under the Fifth Amendment right against self-incrimination. Salinas, a suspect in a murder investigation, was questioned by police without being in custody and without receiving a Miranda warning. Although he answered most of the questions, he remained silent when asked whether shell casings found at the crime scene matched his firearm. At trial, the prosecution introduced his silence as evidence of guilt. Salinas's defense objected, invoking the Fifth Amendment, but the U.S. Supreme Court, in a 5–4 decision, rejected the claim, holding that a suspect must explicitly invoke the right against self-incrimination in order to be protected by it. The Court concluded that the Fifth Amendment does not prohibit the prosecution from using a suspect's silence as evidence of guilt unless the suspect has affirmatively asserted the right. Thus, under this ruling, remaining silent without an explicit invocation of the Fifth Amendment does not shield the suspect from adverse evidentiary use.

For that reason, rules of evidence do not merely obstruct truth; they frequently protect the conditions under which justified belief becomes possible. In the U.S., the elaboration of detailed evidentiary rules is closely connected to the jury institution and to worries that lay decision-makers may overvalue some forms of proof while undervaluing others (Schauer, 2009, p.208). Wigmore (1983), though a rationalist successor to Bentham, recognized that jurors' zeal for truth can heighten vulnerability to prejudice and misinterpretation.[16] Exclusions and limiting doctrines, even when they screen out probative items, can operate as cognitive safeguards that prevent distortion and preserve rational assessment (Schauer, 2009, p. 208).

Eyewitness testimony exemplifies these risks (Loftus and Green, 1980). Once hailed as the *regina probationum*, it is now known to be fragile: perception is fallible, memory is reconstructive, confidence can be socially inflated, and jurors often mistake confidence for accuracy. Unrestricted admission of such testimony may distort rather than clarify judgment. Calibrated corroboration requirements, expert framing of reliability, lineup protocols, and targeted exclusions can therefore *improve* truth-seeking by ensuring that reliance on eyewitnesses is earned by method, not granted by tradition.

A parallel caution applies to "big-evidence" optimism. Digital trails, sensor logs, and high-volume forensic outputs do not automatically raise epistemic quality. Without validated methods, disclosed error rates, and principled aggregation rules that address dependence and confounding, large heterogeneous datasets can manufacture the appearance of convergence. Quantity is not a substitute for warrant; what matters is whether a competent community of reasoners could share, test, and endorse the inferences drawn.

Taken together, these considerations suggest that the relationship between free proof and truth turns on three interlocking conditions: First, evidential base: what actually reaches the tribunal, given party selection, investigative design, and admissibility constraints; Second, standards of appraisal: how the tribunal must reason from that base (standards of proof, burdens, instructions, corroboration and counter-explanation duties), so that conclusions are rationally warranted ex ante rather than merely fortunate *ex post*; Third, epistemic humility, recognition that no accumulation of items guarantees access to objective reality; bounded rationality and institutional design must be accounted for and mitigated. Under these conditions, the free use of evidence can never, by itself, ensure truth. It may raise the opportunity for discovery, but it does not determine the quality of justification. Conversely, well-designed exclusions and reliability screens can raise justificatory quality without abandoning truth as law's telic value.

[16] "The rules of evidence are mainly aimed at guarding the jury from the overweening effect of certain kinds of evidence" (wigmor, 1983, §8a, p. 62).

The practical upshot is an integrated approach that links freedom to responsibility. The use of evidence should be structured so that, while respecting prevailing legal standards (rules of evidence, procedural rights, and legitimacy requirements), the fact-finder is prompted to (a) articulate probative links, (b) confront obvious defeaters (contrary evidence, methodological limits, procedural irregularities), and (c) present reasons in a form that other competent reasoners could recognize as adequate. In short, the operative standard, "the established fact must be sufficiently justified by legally adequate evidence", is best understood as an epistemological task carried out within existing legal norms, not as a license for unbounded inclusion.

On this view, truth remains the aim that tells us what to believe; justification supplies the public reasons that tell us why we are entitled to believe it. Only when evidentiary freedom is disciplined by intersubjective justification does the pursuit of truth become a practice fit for adjudication.

4.2. Strict rules of evidence: a non-determining factor in fact-finding

Critics of exclusionary rules contend that truth is best served by maximal evidentiary openness and by granting fact-finders wide latitude to decide for themselves what to discount. On this view, rules that bar probative material are epistemic obstacles masquerading as fairness. Yet the strictness (or looseness) of admissibility rules is not dispositive of epistemic quality. What ultimately determines the legitimacy of fact-finding is how institutions manage exposure, sequencing, and the intersubjective justification of belief under conditions of bounded cognition. In short, whether a system proclaims "free proof" or "regulated proof," admissibility doctrines alone do not fix how well a tribunal comes to know (Damaška, 2022, p. 199).

Mirjan Damaška helps explain why. He shifts attention away from yes or no questions of admissibility to the psychological and epistemic processes through which evidence is actually evaluated (Damaška, 2022). Modern evidence law, he argues, is evolving beyond authorization and prohibition toward an understanding of "the character of psychological operations involved in evidence processing", (2022, p. 129) fact-finders reason holistically: some items strongly support a proposition, others weakly support or defeat it; and people rarely assign weights monadically and then compute a sum. Wigmore's (1983) analytic chart sought to discipline this complexity by mapping sub-premises and inferential links, "link[ing] a method of reasoning with masses of mixed evidence to a theory of evidence" (Leclerc, Vergès, Vial, 2022). But, as Damaška emphasizes, empirical work shows that it is hard for judges and jurors to disentangle the value of discrete items from global impressions. The aspiration to rational synthesis collides with the realities of human cognition (Damaška, 2022, p. 199).

These cognitive limits become most acute where inadmissible but probative information enters the picture. Damaška's illustration makes this clear: "It is true that

when properly obtained evidence is clearly insufficient for conviction, the law's mandate to exclude illegal but reliable evidence can be effective: judges must acquit. But when properly obtained evidence is compelling, they face a predicament: unless they recuse themselves, they must imagine what a judge uncontaminated by illegal information would decide". Damaška's well-known "hypothetical judge" dilemma captures the problem: if properly obtained evidence is meager, courts can exclude illegally obtained but reliable proof and must acquit. When properly obtained evidence is already compelling, however, exclusion does not erase the mind's prior exposure to the tainted material; the judge must now decide as if she had never learned it. Exclusion preserves procedural integrity but does not, by itself, secure epistemic insulation. Once contamination occurs, the counterfactual basis for justification, what an uncontaminated, trained reasoner would conclude on the admissible record, becomes opaque (Damaška, 2022, p. 129).

A compact hypothetical sharpens the point. Suppose A is found dead. Legally obtained evidence shows that B sent threats to A, was near A's home at the relevant time, and bought a knife matching the murder weapon a week earlier. During an unconstitutional search, police also discover A's bloodstained knife in B's house. The knife is excluded. Even if the judge sincerely resolves to ignore it, exposure to that fact can subtly shape her assessment of the remaining, admissible record - tightening perceived coherence, lowering the salience of defeaters, and raising confidence beyond what the admissible evidence alone would warrant.[17] The verdict might still be accurate, but if the route to it passes through epistemic contamination, its warrant is compromised.

These dynamics challenge the traditional framing of the debate as a contest between strict and liberal regimes. The core question is not which set of admissibility rules is "truer," but how either regime can sustain justified belief given human cognitive architecture. Unchecked truth-seeking can exacerbate bias once illicit yet persuasive information is in play; conversely, exclusion that comes too late, or without shielding, may protect procedure while leaving the epistemic status of the conclusion under-determined. What matters is not whether the final outcome happens to align with reality, but whether the process can be defended as epistemically responsible.

Practically, three levers matter: (a) Sequencing & shielding. Minimize premature exposure to potentially inadmissible material (segmented records, robust suppression hearings, sealed proffers, staggered rulings). Structure the order of proof to reduce salience effects and anchoring; (b) Holistic-but-disciplined evaluation. Use reason-giving templates that map each asserted fact to admissible support, identify defeaters, and explain why those defeaters are neutralized or outweighed, so holistic judgment is constrained by explicit justificatory links rather than impressionistic coherence alone; (c) Intersubjective warrant. Require that (a) an uncontaminated, trained re-

[17] And also, in justifying his decision, he would not explicitly acknowledge that he had been influenced by the illegally obtained evidence. Research related to the influence of illegal evidence can be found in the following source. See Stably, *et al*. (2006*,* p. 469-492).

asoner could reach the same conclusion on the admissible record, and (b) the court can publicly explain that conclusion in terms that are accessible to other competent reasoners. This sets the stage for the standards developed later like intersubjective propositional justification, explanatory coherence, and discursive warrant.

Seen this way, exclusionary rules are not merely moral brakes; they are epistemic safeguards designed to preserve the conditions under which belief can be justified. But they must be embedded in procedures that address exposure and sequencing, or else they risk becoming clean hands over a contaminated mind. The evaluation of evidence is thus not a purely procedural question but an epistemological exercise: the operative standard is that "the established fact must be sufficiently justified by adequate, admissible evidence," where sufficiency is measured against the capacities of bounded reasoners working within an institution that makes reasons shareable and contestable.

This perspective also clarifies the relation to familiar thought experiments. If, in Pardo's scenario, the defendant had testified truthfully, the epistemic problem would not arise, and one might think Laudan's error-minimization program remains intact. The point, however, is not that Laudan is wrong about accuracy, but that rationality must be assessed at the level of publicly shareable and communicatively legitimate justification, not merely at the level of private cognition. Outcomes that are true by accident, or true via contaminated reasoning, fail the justificatory test even if they pass the accuracy test.

Accordingly, justification asks not "Was the result correct?" but "Were we justified in concluding?" It requires cognitively responsible reasoning within evidential and institutional constraints and an epistemically humble posture toward bias and fallibility. Freedom without regulation degenerates into subjective belief; regulation without freedom collapses into mechanical formalism. Truth is the aim; justification is the legitimate path. Legal institutions earn epistemic authority only when their procedures reliably produce publicly defensible reasons for believing established facts.

In sum, even perfect freedom of proof cannot overcome human fallibility. Fact-finding is not a quest for omniscient truth but a practice of constructing justified belief under constraint. Institutions should be judged by how well they prevent or quarantine exposure, discipline holistic reasoning through explicit reason-giving, and anchor verdicts in intersubjective justification. These commitments set the stage for the operative standards that follow, intersubjective propositional justification, explanatory coherence, and discursive norms, as criteria for evaluating the epistemic legitimacy of legal proof.

5. FURTHER DIRECTIONS: FROM INDIVIDUAL RATIONALITY TO INTERSUBJECTIVE JUSTIFICATION

The preceding chapter showed that the opposition between Bentham's evidentiary liberalism, Laudan's epistemic instrumentalism (including his SoP approach), and Damaška's institutional realism cannot be resolved by merely dialing admissibility rules up or down. Each captures a distinct dimension of fact-finding like rational inclusion, error-minimizing design, and structural limits on cognition, yet they converge on a common difficulty: the epistemic status of the facts that courts finally certify. Even if rules were perfectly calibrated to balance truth and fairness, the legitimacy of outcomes would still turn on the justificatory pathway from evidence to conclusion. The familiar Pardo-style scenario already makes the point: a verdict may coincide with the truth while lacking warrant. This does not negate Laudan's truth concerns; it marks the limit of truth as a stand-alone criterion of legitimacy.

Building on this insight, the present chapter shifts from admissibility to the epistemic foundations of legal proof. The guiding question is not which regime best promotes truth in the abstract, but how conclusions can be rationally and publicly justified within adjudication, how institutions can generate reasons that competent others could endorse on the admissible record and that can withstand critical scrutiny. The analysis is comparative in spirit yet agnostic about system-level rankings; the target is the form of justification any system can reasonably demand.

The central move, then, is not to reject Laudan but to extend his program from individual rationality to intersubjective justification. The question is no longer only whether a reasoner could privately minimize error, but whether the conclusion is justified in a way that is (a) reachable by an idealized, uncontaminated, well-trained reasoner on the admissible record, and (b) explainable in public reasons that identify salient defeaters and show why they are neutralized or outweighed. Outcomes that are true by accident, or true via contaminated reasoning, fail this justificatory test even if they pass the accuracy test.

The remainder of the chapter operationalizes these ideas. Section 5.1 articulates Intersubjective Propositional Justification (IPJ) as the baseline: a proposition is justified when it belongs to the decision-maker's evidential base, an idealized trained agent could form the belief on that base, and there are no undefeated, intersubjectively evident defeaters. Section 5.2 adds two complementary standards: explanatory coherence (the verdict must integrate the evidence into a stable, non-ad hoc explanatory structure) and discursive warrant (the court must be able to say why to an appropriate audience, meeting the context-sensitive norm for warranted assertion). Together these standards recast the pursuit of truth from a purely outcome-oriented aspiration into a responsibility-based practice of public justification.

2Legal fact-finding, though epistemically motivated, operates under institutional and communicative constraints that differ from the idealized pursuit of truth in

science or mathematics. The epistemic goal of the legal process is not to achieve certainty but to approximate truth through reasoned deliberation over admissible evidence. To clarify what this entails, it is useful to draw upon the distinction between objective and intersubjective propositional justification, as developed by Silvia De Toffoli (2022) in the philosophy of mathematics. Although De Toffoli's conceptual framework was not designed for legal epistemology, it provides a productive analogy. Just as a mathematical theorem must be justified in a manner comprehensible and verifiable by a community of mathematicians, a proven fact in law must be justified in a manner acceptable to a community of legally competent reasoners. The analogy illuminates a shared structural demand: justification must not only be logically valid but also socially accessible.

According to De Toffoli, objective propositional justification (OPJ) refers to the logical relation between a body of evidence and a proposition it supports.[18] In this framework, justification depends solely on whether the evidence entails or probabilistically supports the proposition, independent of the subject's cognitive limitations or belief state. If all individuals had access to the same evidence and reasoned correctly, they would all share OPJ for the same proposition. However, as De Toffoli (2022, p.246) notes, OPJ alone risks collapsing justification into mere logical entailment, devoid of understanding or explanatory depth. The danger of such objectivism, in her view, is that it strips knowledge of its cognitive and communicative dimensions, it reduces justification to a structural relation rather than a human epistemic achievement.[19] For that reason, she reverses the traditional hierarchy and argues that doxastic justification, justification grounded in a subject's actual belief-forming processes, doxastic justification, should take precedence over propositional justification (De Toffoli, 2022, p. 249).

In contrast, I contend that in the context of legal fact-finding, propositional justification must remain primary. The reasons-first view articulated by Silva and Oliveira supports this position: epistemic reasons, not belief states, are what justify factual conclusions in a public and institutional setting.[20] Legal reasoning, unlike personal conviction, must be articulated in propositional terms that others can evaluate. A fact-finder's justification is legitimate not because it reflects a private mental state but because it can be expressed, examined, and defended in public reason-giving.

To illustrate, consider the case of a jury deliberating on whether a defendant intentionally administered arsenic to a victim. Even if jurors lack specialized chemical

---

[18] "Now consider a case in which I believe some complicated logical theorem T on the basis of sheer guesswork. If T is true, then I have propositional justification to believe that T is true" (Smithies, D. 2015, p. 2783).

[19] De Toffoli indeed agrees with Kornblith in asserting that if propositional justification is regarded as more fundamental than doxastic justification, this leads to skepticism. See Kornblith, (2017, p. 63-80).

[20] "The reasons-first picture characterizes propositional justification in terms of epistemic reasons and doxastic justification in terms of propositional justification" (De Toffoli, 2022, p. 242).

knowledge, they can still possess propositional justification for the claim "arsenic causes death," based on expert testimony or documentary evidence. Their conclusion can be justified even if their personal belief is shallow or incomplete,[21] because what matters in legal epistemology is not the intensity of conviction but the structure of justification. De Toffoli's objection, that justification without understanding is epistemically hollow, raises a legitimate concern. Yet in legal adjudication, the level of understanding required is functional rather than absolute. The law's concern is whether a proposition can be justified to the standard of a reasonable fact-finder, not whether it can withstand ideal philosophical scrutiny. The cognitive depth demanded of the fact-finder is bounded by procedural rationality, not epistemic omniscience.

To bridge the gap between pure objectivity and unbounded subjectivity, De Toffoli introduces the Idealized Capacity Principle (ICP), according to which justification is achieved when an idealized, well-trained human agent could form a justified belief based on the available evidence.[22] This principle has obvious resonance with the legal notion of the reasonable person or reasonable juror. Both invoke an idealized epistemic agent capable of reasoning responsibly within human and institutional limits. Still, ICP faces a further difficulty in complex, adversarial contexts: even idealized agents may disagree on how to interpret competing evidence.

For that reason, De Toffoli (2022, p.258) refines her model into *Intersubjective Propo11sitional Justification* (IPJ), according to which a proposition p is justified for S if and only if (1) the evidence belongs to S's evidential base, (2) an idealized, trained agent would be able to form the belief that p based on that evidence, and (3) the evidence does not admit intersubjectively evident defeaters that remain undefeated. IPJ offers a conceptual key for understanding what it means to prove a fact in law. It articulates a standard that is both epistemic and social: justification requires that the reasons supporting a factual claim be recognizable to others as valid, and that potential defeaters, contrary evidence, counterarguments, or procedural irregularities, be addressed or neutralized. This structure maps directly onto the dynamics of trial deliberation. Judges and juries routinely engage in a collective process of assessing defeaters, weighing competing narratives, and converging toward propositions that can withstand public scrutiny.

At this point, the contrast with Laudan becomes clear. Laudan's model, though epistemically rigorous, remains centered on minimizing error within an individual rational calculus. IPJ, by contrast, relocates epistemic justification into the public domain: the locus of legitimacy is no longer the isolated reasoner but the intersubjective process through which reasons are exchanged, tested, and defended. Intersubjective

[21] See "problem one" (De Toffoli, 2022, p. 249).

[22] "X provides S with propositional justification for p only if an idealized human agent with the appropriate training would likely be in a position to form a doxastically justified belief that p on the basis of X" (De Toffoli, 2022, p. 252).

justification thus preserves Laudan's epistemic ambitions while integrating the procedural and communicative dimensions that his framework leaves implicit.

5.2. Explanatory Coherence and the Discursive Norm of Assertion

If IPJ specifies when a proposition is justified, it leaves open the question of how such justification should manifest in communicative practice. In legal adjudication, justification is not merely held; it must be asserted, explained, and defended. This communicative dimension brings into view the epistemic role of explanation and coherence. Paul Thagard's theory of explanatory coherence offers a framework for understanding this. According to Thagard (1989, as cited in Amaya, 2015), beliefs and hypotheses are justified when they cohere through explanatory relations, when they collectively account for the observed phenomena in a consistent, parsimonious, and mutually supportive way. The more a set of propositions explains and unifies diverse evidence without contradiction, the greater its coherence and, by extension, its justificatory strength.

Amalia Amaya (2015) has applied this model to legal reasoning, arguing that fact-finding in trials operates as a form of coherence-based justification. Legal proof, on this view, consists in weaving together evidence, testimony, and inference into an integrated explanatory structure. Coherence, rather than probability alone, provides the normative metric for justification: it reveals why a conclusion follows from the evidence, not merely that it does. This approach illuminates the epistemic norm underlying the assertion of facts in trials. A judge or juror must be able not only to state a conclusion but also to explain it in a manner consistent with the totality of admissible evidence. Explanatory coherence thus transforms justification from an individual cognitive act into a public, discursive performance. It demands that each asserted fact fit coherently within the broader evidentiary mosaic, forming what Amaya calls the tapestry of reasoning. The strength of this model becomes evident in comparative contexts. When two parties offer competing explanations, one asserting p, the other ¬p, the coherence of each account can be assessed through its capacity to integrate evidence without contradiction or ad hoc supplementation. The more comprehensive and economical the explanation, the higher its epistemic standing. This renders coherence an intersubjectively assessable criterion: it allows participants and observers alike to evaluate the reasonableness of competing factual claims within a shared rational space.

Yet coherence alone cannot account for the social dimension of justification, the fact that legal reasoning unfolds within a communicative forum aimed at persuasion and accountability. For that, Mikkel Gerken's (2012) notion of discursive justification provides a crucial supplement. Gerken argues that assertion, like action, is subject to epistemic norms that depend on context. His Warrant-Assertive Speech Act Principle (WASA) holds that one may assert p appropriately only if one's warrant for believing p is adequate relative to the conversational context (p. 378). In trials,

this means that the epistemic standard for asserting a fact depends on the deliberative context of the courtroom: the presence of adversarial scrutiny, the legal consequences of the claim, and the social expectations of justification. Gerken's framework further refines this through the Discursive Justification-Assertion Principle (DJA), according to which an assertion is justified only if the speaker can articulate the reasons that warrant belief in p to a degree adequate for the discursive context (p.379).

This discursive dimension is integral to legal adjudication. Judges, juries, and lawyers do not merely believe facts, they assert them under institutional conditions that demand reason-giving. A verdict, therefore, is not just a conclusion but a communicative act governed by norms of discursive justification. When a fact is asserted in court, its legitimacy depends not only on its internal coherence (as in Thagard's sense) but also on its capacity to withstand intersubjective evaluation in the communicative space of legal discourse.

Discursive justification also clarifies the distinction between personal belief and institutional assertion. A juror may privately believe a defendant guilty based on intuition, but unless that belief can be discursively justified, explained and defended within the constraints of admissible evidence, it cannot be legitimately asserted in deliberation. Assertion, as Gerken emphasizes, is hearer-directed: it is performed with the expectation that others can understand, evaluate, and potentially accept the reasons offered. This communicative accountability transforms subjective belief into intersubjective epistemic justification. At the same time, being hearer-directed does not mean conforming to audience expectations. An assertion is justified not when it pleases the audience but when it provides reasons that others can rationally endorse.[23] This distinction is vital in legal contexts, where public opinion or political pressure may diverge from epistemic warrant. The integrity of fact-finding depends on maintaining this boundary between persuasion and justification.

When applied to the courtroom, discursive justification reveals how law institutionalizes epistemic responsibility. The trial, as a structured form of public reasoning, compels participants to articulate and defend their factual claims under shared rules of inference and admissibility. These procedural constraints are not impediments to truth-seeking but the very conditions that make intersubjective justification possible. Through them, the epistemic norm of assertion is transformed into a civic practice of accountability.

[23] "Typically, the speaker who makes an assertion has hearer-directed *intentions* in performing a speech act. The speaker may intend the hearer to come to believe something or other about the speaker, or about something else, or intend the hearer to come to desire or intend to do something" (Pagin et Neri, 2021).

5.3. Conclusion: Intersubjectivity as an Epistemic and Civic Norm

Intersubjective propositional justification, supported by explanatory coherence and discursive norms, offers a framework for reconciling truth and fairness in legal proof. It neither abandons Laudan's epistemic ambitions nor returns to Bentham's evidentiary liberalism. Instead, it repositions the epistemic core of adjudication within a communicative and institutional structure of justification. The legitimacy of a verdict, on this view, depends not on the metaphysical truth of its conclusion but on the public reasonability of its justification. Legal fact-finding must therefore be seen as a social epistemic practice, an enterprise of collective reasoning aimed at producing conclusions that are rationally defensible, procedurally valid, and intersubjectively acceptable. Ultimately, this model transforms the pursuit of truth in law into a shared moral and epistemic responsibility. Intersubjective justification ensures that legal reasoning remains both truth-oriented and publicly accountable, sustaining the dual ideals upon which the legitimacy of adjudication depends.

6. CONCLUSION

This article has explored the evolution and limits of truth-seeking in legal adjudication through the comparative analysis of Bentham, Laudan, and Damaška. From Bentham's utilitarian rationalism to Laudan's error-reduction epistemology, both theorists grounded the legitimacy of criminal adjudication in the correspondence between verdicts and reality. Each, in different ways, sought to perfect the fact-finding process by aligning it more closely with truth. Yet their shared rationalist optimism, faith in the capacity of human reason to reach objective accuracy through evidentiary freedom, proves epistemically incomplete.

The assumption that free proof naturally conduces to truth neglects the bounded and situated character of human judgment. Fact-finders do not process information as neutral observers but as cognitively limited agents operating within institutional frameworks. Adjudication, therefore, is not merely a logical reconstruction of past events; it is a structured epistemic practice conditioned by procedural rules, moral commitments, and cognitive constraints. As Damaška's institutional analysis reveals, the architecture of adjudication, the distribution of roles, evidentiary thresholds, and procedural safeguards, fundamentally shapes how facts are constructed and how belief is justified. Truth, in this context, is not simply discovered but institutionally mediated.

Accordingly, this paper has argued that the legitimacy of adjudication cannot rest on accuracy or error minimization alone. It depends on epistemic justification, the disciplined process by which factual beliefs are warranted under human and institutional limits. To know truly in law is not merely to reach the correct result, but to reach it through reasons that are rational, transparent, and publicly defensible. A

verdict that happens to be true but lacks epistemic warrant is normatively deficient; it confuses truth with legitimacy and fact with knowledge.

From this perspective, exclusionary and procedural rules acquire renewed meaning. They are not external constraints on truth-seeking but internal conditions for epistemic responsibility. By disciplining the process of reasoning and ensuring fairness, they transform fact-finding from an exercise in persuasion into an exercise in justification. Legal conclusions must therefore be sufficiently justified by admissible evidence, and that justification must be articulated in forms that are intelligible and contestable within the legal community.

Intersubjective justification thus emerges as the cornerstone of legitimate fact-finding. Legal decisions must be capable of surviving critical scrutiny and rational disagreement, not because they are infallible, but because they are reasonably warranted under shared epistemic and procedural standards. This intersubjective dimension anchors adjudication in a community of reason, ensuring that verdicts rest on communicable and contestable grounds rather than on individual conviction.

Ultimately, this paper proposes a reorientation of evidentiary philosophy: from truth as outcome to justification as process. The law's epistemic ambition should not be omniscient truth, but justified belief within bounded rationality. A legal system committed to epistemically justified fact-finding, one that acknowledges cognitive limitations, values procedural discipline, and institutionalizes shared reasoning, best reconciles the dual aims of justice and truth. Only through such a framework can adjudication sustain both its moral authority and its epistemic integrity.

REFERENCES

Amaya, A. (2015). *The Tapestry of Reason: An Inquiry into the Nature of Coherence and its Role in Legal Argument*. Bloomsbury.

Beccaria, C. (1995). *On Crimes and Punishments and Other Writings*. Cambridge University Press.

Bentham, J. (1827). Rationale of Judicial Evidence (Vol. 1). Hunt and Clarke.

Bentham, J. (1843). *Rationale of Judicial Evidence*. In J. Bowring (Ed.), *The Works of Jeremy Bentham* (Vol. 7, pp. 196-453). William Tait.

Bentham, J. (1907). *An Introduction to the Principles of Morals and Legislation*. Clarendon Press.

Busey, T., & Loftus, G. (2007). Cognitive science and the law. *Trends in Cognitive Sciences*, *11*(3), pp. 111-117.

De Champs, E. (2015). *Enlightenment and Utility: Bentham in French, Bentham in France*. Cambridge University Press.

De Toffoli, S. (2022). Intersubjective propositional justification. In P. Silva Jr. & L. R. G. Oliveira (Eds.), *Propositional and Doxastic Justification: New Essays on Their Nature and Significance* (pp. 241-262). Routledge.

Damaška, M. (1997). *Evidence Law Adrift.* Yale University Press.

Damaška, M. (2019). *Evaluation of Evidence: Premodern and Modern Approaches*. Cambridge University Press.

Damaška, M. (2022). The Future Of Free Proof In Criminal Cases. In K. Bárd (Ed.), *Liber Amicorum Károly Bárd II: Constraints on Government and Criminal Justice* (pp. 199-207). L'Harmattan.

Gerken, M. (2012). Discursive justification and skepticism. *Synthese*, *178*, pp. 373-394.
Guthrie, C., Rachlinski, J. J., & Wistrich, A. J. (2001). Inside the judicial mind. *Cornell Law Review*, 86(4), pp. 777-830.
Haack, S. (2014). *Evidence Matters: Science, Proof, and Truth in the Law*. Cambridge University Press.
Kahneman, D. (2011). *Thinking, Fast and Slow.* Farrar, Straus and Giroux.
Kornblith, H. (2017). Doxastic justification is fundamental. *Philosophical Topics, 45(1),* pp. 63-80.
Laudan, L. (2006). *Truth, Error, and Criminal Law: An Essay in Legal Epistemology.* Cambridge University Press.
Leclerc, O., Vergès, É., & Vial, G. (2022). Turning a graphical method of evidential reasoning into an operational tool for judges? Empirical evidence. *International Journal of Evidence and Proof*, *26*(2), pp. 136-160.
Loftus, E. F., & Green, E. (1980). Warning: Even memory for faces may be contagious. *Law and Human Behavior, 4*(4), pp. 323-334.
Neilson, G. (1890). *Trial by Combat.* Williams and Norgate.
Pagin, P., & Neri, M. (2021). Assertion. In E. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy.*
Pardo, M. (2005). The field of evidence and the field of knowledge. *Law and Philosophy, 24,* pp. 321-379.
Schauer, F. (2009). *Thinking Like a Lawyer: A New Introduction to Legal Reasoning.* Harvard University Press.
Smithies, D. (2015). Ideal rationality and logical omniscience. *Synthese, 192*(9), pp. 2783-2811.
Stably, N., et al. (2006). The impact on juror verdicts of judicial instruction to disregard inadmissible evidence: A meta-analysis. *Law and Human Behavior*, *30*, pp. 469-492.
Tachau, K. (1988). *Vision and Certitude in the Age of Ockham: Optics, Epistemology and the Foundations of Semantics 1250–1345*. E. J. Brill.
Twining, W. (1985). *Theories of Evidence: Bentham and Wigmore*. Stanford University Press.
Twining, W. (2006). *Rethinking Evidence: Exploratory Essays* (2nd ed.). Cambridge University Press.
Twining, W. (2019). Bentham's theory of evidence: Setting a context. *Journal of Bentham Studies*, *18(1),* pp. 27-39.
Wigmore, J. H. (1983). *Evidence in Trials at Common Law* (Vol. 1, P. Tillers, Ed.). Little, Brown & Co.

Jurisprudence and legislation

Griffin v. California, 380 U.S. 609 (1965).
Salinas v. Texas, 570 U.S. 178 (2013).
Tehan v. United States, 383 U.S. 406 (1966).

Quaestio facti. Revista Internacional sobre Razonamiento Probatorio
Quaestio facti. International Journal on Evidential Legal Reasoning
Sección: Homenaje a Larry Laudan
2026 I 10 pp. 187-219
Madrid, 2026
DOI: 10.33115/udg_bib/qf.i10.23222
Marcial Pons Ediciones Jurídicas y Sociales
© Michael S. Pardo
ISSN: 2604-6202
Recibido: 17/10/2025 | Aceptado: 15/12/2025 | Publicado *online*: 19/01/2026
Editado bajo licencia Reconocimiento 4.0 Internacional de Creative Commons

LEGAL EPISTEMOLOGY AND LEGAL PROOF

Michael S. Pardo*

ABSTRACT: This article examines the relationship between legal epistemology and legal proof. On the one hand, there has been an explosion of recent scholarship applying philosophical ideas, concepts, and arguments from epistemology to the law of evidence and the process of proving facts in legal settings. On the other hand, there has been ongoing theoretical debates about the law itself, focusing on the best understanding of the evidentiary proof process and its component parts (including the law of evidence). These are distinct, but related, theoretical projects.

The article begins by presenting a picture of legal proof in the United States. The discussion focuses on details of the evidentiary proof process that have been relatively neglected in legal epistemology. These details include the explanatory structure of the proof process; the role of the parties in the adversarial system; and how courts evaluate the sufficiency of evidence. In light of the picture of legal proof presented, the article then examines legal epistemology by the philosopher Larry Laudan. A number of lessons emerge from this discussion for applying epistemology to evidence law and legal proof. Finally, the article applies these lessons to other issues in recent legal epistemology, including: statistical evidence; the "preponderance of the evidence" standard; epistemic safety; knowledge; and epistemic injustice.

* Scott K. Ginsburg Professor of Evidence Law, Georgetown University Law Center. This article is based on my plenary lecture at the 2nd Michele Taruffo Girona Evidence Week in June 2025. I was fortunate to meet Michele Taruffo at a conference in Girona in 2011 and to have engaged with him in print in a 2019 symposium. I was honored to give a lecture at the conference bearing his name. My thanks to Jordi Ferrer Beltran and Diego Dei Vecchi for inviting me to deliver the lecture and to the conference participants for many helpful comments and questions. My thanks also to Ron Allen, Meredith Render, and two anonymous referees for helpful comments on previous drafts. Both this article and my lecture are dedicated to the memory of Larry Laudan—whom I am grateful to have known and to have learned so much from.

KEYWORDS: Legal proof, standards of proof, sufficiency of evidence, probative value, legal epistemology, Larry Laudan, inference to the best explanation, error allocation, error reduction, statistical evidence, preponderance of evidence, safety, knowledge, epistemic injustice

SUMMARY: 1. INTRODUCTION.— 2. A PICTURE OF LEGAL PROOF.— 3. THE LEGAL EPISTEMOLOGY OF LARRY LAUDAN: 3.1 Error allocation and error reduction; 3.2 The illusion of precision and the explanatory structure of proof.— 4. CONTEMPORARY LEGAL EPISTEMOLOGY: 4.1 Statistical evidence; 4.2 The "preponderance of the evidence" standard; 4.3 Epistemic safety, 4.4 Knowledge; 4.5 Epistemic injustice.— 5. CONCLUSION.— BIBLIOGRAPHY

1. INTRODUCTION

This article examines the relationship between legal epistemology and legal proof. On the one hand, there has been an explosion of recent scholarship applying philosophical ideas, concepts, and arguments from epistemology to the law of evidence and the process of proving facts in legal settings. On the other hand, there has been ongoing theoretical debates about the law itself, including the best understanding of the evidentiary proof process and its component parts (including the law of evidence). It is important to recognize these as distinct, but related, theoretical scholarly projects. By "legal epistemology," I mean the use of philosophical concepts and ideas from epistemology—for example, knowledge, justification, warrant, safety, sensitivity, and epistemic injustice—to illuminate aspects of legal evidence and proof. By "legal proof," I mean the formal process, regulated by the law of evidence, by which contested or disputed legal facts are considered to be "proven" for legal purposes.

The discussion will be broad in scope. I will try to sketch out some of the possibilities for, and some of the limitations on, using epistemology to understand the law. Work in legal epistemology by the philosopher Larry Laudan will be used as a detailed example to ground the analysis. Then, the discussion will apply lessons from the analysis of Laudan's scholarship to other recent work in legal epistemology. But first, the article will present a picture of legal proof and some of the possible ways in which epistemology may illuminate the evidentiary proof process.

Three caveats are necessary. First, the discussion is not meant to be a comprehensive overview of the ways in which epistemology can contribute to understanding legal proof. Rather, the article aims to present one particular vision in which "legal epistemology" is understood as continuous with "evidence theory" (or theoretical accounts of evidence law) more generally.[1] In this vision, epistemology brings an ad-

[1] This is a very loose reference to the philosopher W.V. Quine's notion of philosophy as "continuous with science" (Quine, 1969, p. 126). The only similarity to the Quinean project that I wish to emphasize is the following. In the vision presented, legal epistemologists are essentially in the same boat as evidence scholars generally in trying to make sense of the part of reality dealing with law's evidentiary proof practices—with the former focused on some of the more abstract aspects of that reality and brin-

ditional set of conceptual tools and ideas—alongside the tools from other academic disciplines—that may help us to better understand legal proof (and thus also perhaps to critique and improve it).

Second, the focus in this article is on ways in which epistemology may contribute to understanding law, not the other way around. A potentially apt methodological notion from the philosopher W.V. Quine is the idea of "reciprocal containment" (Quine, 1969, p. 83). Quine was referring to the relationship between epistemology and psychology in his image of "naturalized epistemology"—with epistemology containing and making use of psychology and other empirical information, and psychology simultaneously containing and studying aspects of epistemology. I am suggesting a similar kind of "reciprocal containment" between law and epistemology—with law making use of epistemology to the extent that it, like other tools, helps to illuminate law, and with epistemology studying law as an example of an important institutionalized epistemic context, among other social epistemic practices. A corollary of this relationship is that there may be reasons *internal* to philosophy for philosophers to discuss law and legal examples that do not necessarily shed much, if any, light on legal proof. But that is not the focus of this article. The focus is on what use (pragmatic or theoretical) epistemology may have for law.

Third, the discussion will focus on evidence and legal proof in the United States.[2] I am not a legal comparativist, and the discussion will not attempt to compare proof practices among different legal systems. Nevertheless, the article aims to reveal some general lessons through a sustained focus on legal proof in the United States. Indeed, as the article will emphasize, part of the challenge—and a potential limitation with applying epistemology to law—is the risk of mismodeling or misdescribing the complex practices, rules, and doctrine that make up legal proof in particular jurisdictions. To illustrate this, I will often use as examples standards of proof and the epistemic sufficiency of evidence in the United States.

The article proceeds as follows. Part 2 spells out a general picture of legal proof in the United States. Part 3 discusses the legal epistemology of Larry Laudan, with a particular focus on standards of proof. Part 4 applies the lessons from the analysis in Part 3 to recent work in legal epistemology. This Part discusses five examples: statistical evidence, the "preponderance of the evidence" standard, epistemic safety, knowledge, and epistemic injustice. Part 5 concludes with some general reflections on the relationship between legal epistemology and legal proof.

---

ging to bear a set of philosophical tools and concepts. But, in the vision presented, there is no clear demarcation between the two enterprises (legal epistemology and evidence theory), nor does epistemology necessarily provide a privileged *a priori* foundation or vantage point by which to assess those practices.

[2] This is itself somewhat of an over-simplification, as there is some variation in the evidence law and proof practices among the different State jurisdictions in the United States. My focus will be on legal proof at the federal level and the constitutional issues that apply generally throughout the United States.

2. A PICTURE OF LEGAL PROOF

By "legal proof," I mean the formal process, regulated by the law of evidence, by which contested legal facts are "proven" (or not) for legal purposes.[3] The formal process of legal proof has many component parts. These include the practices of various actors as well as official rules and legal doctrines. These practices, rules, and doctrines vary across jurisdictions. In order to reduce some of this complexity and variance, I will start by introducing a simple picture of legal proof: *The presentation of evidence, on one hand, in light of a burden and standard of proof, on the other*. In this simple picture, a fact has been proven when the "probative value" (or strength or quality) of the evidence satisfies or surpasses the applicable standard of proof, and an alleged fact has not been proven when it falls short of the standard.[4]

In the United States, there are three common standards of proof: "preponderance of the evidence" (POE), "clear and convincing evidence" (CCE), and "beyond a reasonable doubt" (BARD).[5] Although the standards are employed with relatively vague, open-ended, and sometimes ambiguous language, they are intended to impose increasingly higher burdens on the party with the burden of proof. Among the three, BARD is the highest, POE is the lowest, and CCE is an intermediate standard.[6] As matter of legal doctrine, the standards apply to each of the *elements* of a crime, civil cause of action, or an affirmative defense. The basic policy idea behind the standards of proof is that the preponderance standard aims to treat the parties roughly equally with regard to the risk of error[7]—by favoring whichever side the evi-

---

[3] There is a sense of the word "prove" that is factive. In other words, in this sense, something can be "proven" only if it is true—a false proposition cannot be "proven." To be clear, this is *not* the sense in which I am using "prove" in the context of legal proof. Rather, a disputed factual proposition is "proven" for legal purposes when a party has satisfied a *burden of proof*, and a fact is not "proven" when they fail to do so. When a party has satisfied a burden of proof, the fact is considered "proven" for subsequent legal purposes (*i.e.*, it is treated *as true* by the law for purposes of attaching legal consequences). But propositions "proven" within the process of legal proof might, in fact, be false.

[4] In the United States, "probative value" is an important concept for both admissibility and sufficiency of evidence. See Federal Rule of Evidence 403 (giving the trial judged discretion to exclude an item of evidence when its probative value is substantially outweighed by one or more "dangers" such as unfair prejudice, confusing the issues, misleading the jury, or efficiency considerations). Whether evidence is sufficient or not to meet a standard of proof depends on what a "reasonable jury" could conclude about the probative value of the evidence as a whole. See Pardo (2023a); Pardo (2013).

[5] Standards of proof specify the standard of *persuasion* for the party with the burden of proof on the elements of a civil cause of action, a crime, or an affirmative defense. See Colorado v. New Mexico, 467 U.S. 310, 315 (1984); Addington v. Texas, 441 U.S. 418, 423 (1979) ("the evolution of this area of law has produced across a continuum three standards or levels of proof for different types of cases.").

[6] *Ibid*.

[7] See E.M.D. Sales, Inc. v. Carrera, 604 U.S. 45, 50 (2025) ("the preponderance-of-the-evidence standard has remained the default standard of proof in American civil litigation. That default makes sense: The preponderance standard allows both parties in the mine-run civil case to share the risk of error in roughly equal fashion.")

dence appears to support—with higher standards such as CC and BARD attempting to skew the risk of error in a particular direction.[8] Thus, in the simple picture of legal proof: *the law asks a fact-finder (judge or jury) to decide, based on the admissible evidence, whether the disputed fact (a legal element) is proven to the applicable standard of proof (POE, CCE, or BARD).*

Epistemology may help to illuminate different aspects of this picture. It may reveal something about qualities of the evidence itself—for example, its relevance or probative value. Or it may illuminate aspects of the standards of proof and what they require. Or it may help illustrate how the evidence and the standard of proof fit together and the reasoning that connects them. In thinking about the different possible ways in which epistemology may apply to legal proof, a helpful distinction is one between *micro-level* and *macro-level* analysis of the proof process.[9] Micro-level analysis focuses on individual items of evidence and may be useful for answering questions about relevance, probative value, or the application of various admissibility rules. Macro-level analysis focuses on cases as a whole and may be useful for answering questions about burdens and standards of proof, the sufficiency of evidence, and the application of evidentiary presumptions.[10]

In the abstract, epistemology is potentially a source of illumination or understanding for legal proof for the simple reason that legal fact-finding could be epistemically better or worse. There are different kinds of decision-making errors that could be made in the legal proof process—examples include: errors in assessing the relevance or probative value of items of evidence; errors in concluding whether or not the evidence as a whole meets a standard of proof; and, ultimately, errors in whether the facts that have been "proven" (or not) for legal purposes, in fact, match or correspond with reality.[11] But moving from abstract possibilities to particular legal systems requires a detailed understanding of the legal reality. In other words, in order to understand how epistemic issues arise in law, it is important to fill in additional details to our simple picture, if the analysis is intended to shed light on or help us to understand the law.

As an example, I will focus on the issue of the sufficiency of evidence. In the United States, an important, recurring issue in both civil and criminal cases is whether the evidence is sufficient to support a contested factual finding.[12] In assessing the suffi-

---

[8] See Cruzan v. Missouri Dep't of Health, 497 U.S. 261, 283 (1990) ("The more stringent the burden of proof a party must bear, the more that party bears the risk of an erroneous decision."); In re Winship, 397 U.S. 358, 371 (1970) (Harlan, J., concurring).

[9] See Pardo (2013, p. 562-69).

[10] There is some porousness in the distinction between the micro- and macro-levels. For example, an evidentiary presumption might apply to both individual items of evidence at the micro-level and also might affect the sufficiency of evidence at the macro-level.

[11] See note 3, *supra*.

[12] One exception is acquittal in a criminal case—which is not subject to sufficiency-of-evidence review in order to protect the jury's constitutional power to nullify (or to refuse to convict despite the evidence meeting the standard of proof).

ciency of evidence, as a matter of legal doctrine, courts in the United States generally ask whether a "reasonable" or "rational" jury could (or must), based on the evidence, find the disputed fact to the standard of proof.[13] For example, could a reasonable jury in a civil contract lawsuit find for the plaintiff by a POE? Or in a criminal case, could a reasonable jury find for the prosecution BARD? If a reasonable jury could not find for a party—and therefore must find for the opponent, given the evidence and the standard of proof—that will generally end the case.[14] Sufficiency assessments of the evidence by judges are critically important pre-trial, during trial, and post-trial. These rulings may decide whether a case gets to trial in the first place; whether a trial goes to a jury for decision; or whether a jury verdict is allowed to stand or is overturned as a matter of law.[15]

Decisions by judges on the sufficiency-of-evidence question are essentially *epistemic* determinations (as are fact-finding decisions at trial). Sometimes the evidence is strong enough, and sometimes not, given the applicable standard of proof. Importantly, sufficiency determinations are also decisions about whether a contested fact has been "proven" or not for legal purposes—and thus these determinations are an integral part of legal proof.[16]

In order to understand the epistemic sufficiency of evidence in the United States, it is also necessary to understand two related aspects that structure the proof process. These aspects are (1) the adversarial system, and (2) the fact that the proof process is structured around competing possible explanations of the disputed events and the evidence. In addition to relying on the parties to present (and object to) evidence, according to the "party presentation" and "master of the case" principles, parties provide explanations that purport to explain disputed events in light of the evidence.[17] The parties use evidence and arguments to support or challenge these explanations, and fact-finders evaluate and select among these explanations (although they are permitted, but not required, to construct their own explanations) (Pardo & Allen, 2008; Amaya, 2019). According to the "relative plausibility" account of legal proof that Ron Allen and I have defended: as a descriptive matter, we can best understand

---

[13] See Jackson v. Virgina, 443 U.S. 307, 319 (1979); Reeves v. Sanderson, 530 U.S. 133, 149-53 (2000).

[14] See Federal Rule of Civil Procedure 56 (authorizing motions for summary judgment based on insufficient evidence); Federal Rule of Criminal Procedure 29 (authorizing motions for judgment of acquittal based on insufficient evidence).

[15] For example, a court may grant summary judgment to a party pre-trial, see Federal Rule of Civil Procedure 56; a court may grant a judgment as a matter of law to a party at the close of evidence (and thus not send the case to the jury), see ibid at Rule 50; or a court may grant a judgment as matter of law after a trial (despite a jury verdict to the contrary), see *ibid*.

[16] See note 3, *supra*.

[17] See United States v. Sineneng-Smith, 140 S. Ct. 1575, 1579 (2020) ("In our adversarial system of adjudication, we follow the principle of party presentation."); Greenlaw v. United States, 554 U.S. 237, 243 (2008) ("in both civil and criminal cases … we rely on the parties to frame the issues for decision and assign to courts the role of neutral arbiter of matters the parties present."). For further discussion of these principles, see Allen, Pardo, Lawrence & Smiciklas (2025, p. 283-92).

the standards of proof as providing explanatory thresholds or explanatory burdens within this process (Allen & Pardo, 2019a; Allen & Pardo, 2019b). The preponderance standard requires having the better or the best of the competing possible explanations. Higher standards impose higher explanatory thresholds—for example, a plausible defense explanation may raise a reasonable doubt even if it is not better than the prosecution's explanation.

Similar to fact-finding at trial, courts evaluate the sufficiency of evidence in light of explanatory considerations (Pardo, 2023a). For example, in a civil case under the preponderance standard, a court evaluates whether a "reasonable jury" could find the plaintiff's explanation better than the alternatives based on the evidence. In making sufficiency determinations, courts refer to explanatory relationships between the evidence and the explanations of the parties. These include (1) consistency and inconsistency between evidence and explanations; (2) the absence of evidence to support important parts of explanations; (3) counterfactual considerations; (4) fit with background knowledge; and (5) the absence of plausible alternatives (Pardo, 2023a, p. 453-54). The epistemic evaluation in sufficiency determinations falls in between two extremes. On one hand, it is not purely *objective* in the sense that courts are discovering objective conditional probabilities (i.e., the "objective" probability of the disputed fact, given the evidence). On the other hand, the process is not purely *subjective* either, in the sense that all decisions are considered to be "reasonable" or "rational" no matter how weak the evidence. The former is a type of epistemic fantasy and the latter is a type of epistemic nightmare. The reality of sufficiency review is an epistemic practice between these extremes.

For the discussion to follow, it is important to notice three important details about the picture of legal proof spelled out in this Part:

— First, the threshold for epistemic sufficiency is a *minimal* one.[18] Courts do not attempt to discern the optimal or maximally rational decision in light of the evidence. Courts ask whether a "reasonable" jury *could* find for the party with the burden of proof on that issue.[19]

— Second, in general the epistemic context is a *permissive* one. In cases that proceed to a jury verdict following a trial, either outcome is epistemically permissible (the evidence is "sufficient" to support either finding).[20]

---

[18] See Allen, Pardo, Lawrence & Smiciklas (2025) for further discussion of how "minimal rationality" best explains the law of evidence in the United States.

[19] See note 13, *supra*.

[20] See notes 12-15, *supra*. Smith (2018, p. 1205-06) argues that there is something problematic about two courts reaching different verdicts based on similar evidence ("If two courts were presented with equivalent bodies of evidence against two individuals charged with equivalent crimes, could it really be acceptable for them to reach *different* verdicts … there seems to be something *viscerally* bad about such a turn of events"). But, within the United States, the permissibility of either verdict is what allows for the case to proceed to trial in the first place (or go to a jury following a trial). If only one verdict were permissible based on the evidence, then this would be a reason to end the case as a matter

— Third, the process is highly dependent on the adversarial system. The parties are responsible for producing evidence and for providing the competing explanations that will structure the proof process. Judges and juries can decide based on explanations not advanced by the parties (but they are not required to).

These are facts about legal proof in the United States and how it is structured. This Part spells them out in detail because, in my view, they have been neglected in recent epistemological discussions of legal proof—which have focused primarily on fact-finding at trial and in determining the epistemically correct outcome. Focusing solely on the latter issues, however, misses the broader context for when evidence is epistemically sufficient and, thus, when facts are "proven" for legal purposes. Epistemology may, of course, provide reasons to critique the features or practices that I have spelled out. And nothing in the analysis to follow assumes that the picture of legal proof presented, or any of its constituent features, cannot be put into question by epistemological analysis. But when epistemology is being used to try to illuminate or help us understand aspects of legal proof—or when normative or prescriptive epistemological analysis relies on assumptions about features of legal proof—there is a risk of mismodeling the object of inquiry by missing this broader context.

Parts 3 and 4 now turn from this picture of legal proof to legal epistemology. The article next discusses work by Larry Laudan and then turns to other examples in recent legal epistemology. The aim is to illustrate both positive ways in which epistemology can illuminate legal proof as well as some limitations on using epistemology to understand aspects of legal proof.

3. THE LEGAL EPISTEMOLOGY OF LARRY LAUDAN

After a distinguished career as a philosopher of science,[21] Larry Laudan later in his life turned to law and legal proof—with a particular focus on criminal law and standards of proof in the United States. Despite the implicit connections between evidence law, legal proof, and epistemology, when Laudan turned his attention to law, modern evidence scholarship had not made extensive use of modern work in epistemology. There were, of course, exceptions, going back at least to Jeremy Betham (1843) who wrote that, "the field of evidence is no other than the field of knowledge."[22] But in his 2006 book, *Truth, Error, and Criminal Law: An Essay in*

---

of law. See notes 13-14, *supra*. The only reason for a criminal trial to proceed to a jury verdict in such a circumstance would be to preserve the jury's power to acquit even if the evidence appears to satisfy the standard of proof. If, on the other hand, the evidence is insufficient, and acquittal is the only permissible outcome, then the case should not proceed.

[21] For an excellent overview of Laudan's many contributions to the philosophy of science, see Psillos (2023).

[22] See also Twining (2019).

Legal Epistemology, Laudan was able to write, plausibly, that "[l]egal epistemology … scarcely exists as a recognized field of inquiry" (p. 2).[23]

Laudan set out to change this state of affairs. And he was largely successful in this endeavor. In my view, he was partly responsible for the explosion in scholarly work on legal epistemology in recent years.[24] In her biography on HLA Hart, Nicola Lacey (2004, p. 151) quotes Hart as describing his own work as, "sell[ing] just a little philosophy to the lawyers."[25] But, as John Mikhail (2007, p. 763-64) has argued, Hart's legacy also includes "selling law to the philosophers" and "thereby transforming the discipline of analytic philosophy itself."[26] We can say something similar about Larry Laudan and the field of legal epistemology: he both sold some epistemology to the evidence scholars, and he sold some evidence law to the philosophers. And the explosion of work in legal epistemology in the past couple of decades by both evidence scholars and by philosophers is a testament to his legacy. It is thus no longer true, as Laudan wrote in 2006, that "[l]egal epistemology scarcely exists as a recognized field of inquiry" (p. 2).[27]

The discussion in this Part will first mention some of Laudan's insights in legal epistemology. Then, it will focus on some limitations in his analysis for understanding legal proof, based on the picture of legal proof outlined in Part 2. (Part 4 will illustrate how the critique and similar limitations arise in other recent work in legal epistemology.)

Important insights from Laudan's work include the following:

— Laudan argued that legal standards of proof should focus less on the subjective state of mind of the fact-finder (such as doubt, belief, certainty) and should focus more on the features of the evidence that would justify or warrant someone

---

[23] Laudan (2006, p. 2) viewed his work on law as "applied epistemology" or "the study of whether systems of investigation that purport to be seeking the truth are well engineered to lead to true beliefs about the world." For a different conception of legal epistemology—one that sees epistemic properties in possible tradeoffs with accuracy—see Enoch, Fisher, & Spectre (2021).

[24] Laudan was helped in this endeavor by several other philosophers, including: Alvin Goldman, Judith Thomson, Jonathan Cohen, Susan Haack, Brian Leiter, Fred Schauer, and Scott Brewer, among many others. But in my view Laudan's work was critical in spurring interest in legal epistemology. In addition to his scholarship, Laudan also helped to generate interest in legal epistemology by hosting conferences in Mexico City that brought together legal scholars and philosophers from around the world. I was fortunate to be able to participate in two of them (this is where I first met Larry in person).

[25] Lacey (2004, p. 155-78) describes the methodological shift toward analytic philosophy that Hart helped to bring about for legal theory.

[26] See also Mikhail (2007, p. 764) ("Drawing on his background as a lawyer, Hart imported a much-needed practical sensibility into [ordinary language philosophy] at a time that it risked becoming stale and complacent. He also exerted a powerful influence on his students and colleagues, including Austin, which has not always been fully appreciated.").

[27] See note 23, *supra*. See also Enoch, Fisher, & Spectre (2021, p. 85) ("Legal epistemology seems to be exploding. More and more philosophers seem to be taking an interest in the theory of evidence law, and to bring along with them to legal theory the freshest news from the abstract study of epistemology").

in holding a particular state of mind.[28] This analysis usefully drew attention to the epistemic dimensions of legal proof.[29]

— In theorizing about errors in legal proof, Laudan explained and illustrated an important distinction between: *probatory* and *material* errors (Laudan, 2006, p.12). A fact-finder might error in applying the standard of proof (for example, finding a defendant guilty BARD when the evidence does not warrant that conclusion) or a fact-finder might error by making a finding that does not match reality. The former is a *probatory* error and the latter is a *material* error. Laudan clarified that neither type of error implies the other (and that evidence theory often confuses the two).[30]

— Laudan clarified and illuminated the presumption of innocence (Laudan, 2016, p. 90-109). He argued, persuasively in my opinion, that the presumption is best understood as a *probatory* epistemic blank slate, and not as itself evidence or as a presumption about material innocence.

— Laudan's work (some co-authored with Ron Allen) usefully illustrated some of the costs associated with false acquittals in criminal cases as well as the conceptual limits of Blackstone ratios when justifying standards of proof.[31]

My critique will focus on two related aspects of Laudan's analysis of standards of proof in the United States (the legal system that was the primary focus of his legal epistemology). The first aspect concerns Laudan's attempt to isolate the policy justifications and goals for standards of proof. The second aspect concerns an assumption that standards of proof (or their outcomes) can be specified with precision. Each aspect is discussed in turn.

3.1. Error allocation and error reduction

Laudan's analysis of standards of proof attempted to reduce and isolate features of the law that cannot be reduced and separated from the proof process as a whole. To this end, Laudan argued in favor of two related claims (1) that the sole function of standards of proof should be to allocate or distribute fact-finding errors between the parties, and (2) that no other rules or practices should serve this function.[32]

---

[28] See Laudan (2006, p. 31) ("The most earnest jury, packed with twelve people desirous of doing the right thing and eager to see that justice is done, are left dangling with respect to how powerful a case is required before they are entitled to affirm that they believe the guilt of the defendant beyond a reasonable doubt").

[29] Standards of proof have both psychological and epistemological aspects—the former concerns if and when the fact-finder is persuaded that the standard is met; the latter concerns if and when the fact-finder ought to be persuaded.

[30] *Ibid*. In other words, there can be a probatory error without a material error, and there can be a material error without a probatory error.

[31] See Laudan (2016); Allen & Laudan (2008).

[32] This was a consistent theme is his analysis. See Laudan (2006, p. 76); Laudan (2015, p. 4); Laudan (2016, p. 110-37).

In defending the first claim, Laudan asserted that: "the standard of proof is not … an error *reducing* device" and that it functions solely to "shift the errors that do occur in a direction that reflects their respective costs" (Laudan, 2015, p. 4 [emphasis added]).[33] As a conceptual or analytical matter, this is incorrect. In addition to *allocating* the risk of error, standards of proof may also affect the total number of errors a system creates. For example, the preponderance standard is thought, under certain conditions, to *reduce* total errors.[34] Therefore, raising a standard of proof above the preponderance standard may, as a consequence, produce more overall errors.[35] In addition to error allocation, error reduction is thus also an important policy goal when considering standards of proof.[36]

In defending the second claim, Laudan argued that all error *allocation* should be funneled into the standard of proof. Therefore, he argued, all other evidence rules or practices that had some asymmetric affect (*i.e.*, systemically favoring one side or the other) should be eliminated. He referred to rules and practices having this effect as "burden enhancers" and "burden reducers" and called for their elimination from the process of proof.[37] This is an interesting and illuminating philosophical exercise, but it is a kind of epistemic fantasy.[38] It would be virtually impossible to eliminate every possible rule or practice that might systematically help one side more than the other. And in the United States, which was Laudan's primary focus, this is a non-starter because many of the burden-enhancing rules and practices are constitutional rights for criminal defendants under the Fourth, Fifth, and Sixth Amendments of the United States Constitution. There is an elegance to Laudan's analysis on this issue, but it is more realistic—and more helpful from an evidence-scholar perspective—to focus on the risk of error in the system as whole, with standards of proof as one component.[39] This is because of the second aspect of the analysis discussed below.

---

[33] Laudan argued that other rules (such as admissibility rules) should focus on the goal of error reduction. Ibid.

[34] See United States v. Fatico, 458 F.3d 388, 403 (E.D.N.Y. 1978) (Weinstein, J.) ("As a general rule, a 'preponderance of the evidence' standard is relied upon in civil suits where the law seeks … to minimize the probability of error"); see also Cheng & Pardo (2015).

[35] This follows from the fact that some factual findings will be more likely true, given the evidence, but will be taken to be false (or not "proven") by the law.

[36] This is why the United States Supreme Court has adopted the POE standard for the application of admissibility rules in the *Federal Rules of Evidence*, in both civil and criminal cases. See Bourjaily v. United States 483 U.S. 171 (1987); Huddleston v. United States, 485 U.S. 681 (1988). See also Pardo (2023b)

[37] For a detailed list and discussion of "burden enhancers" and "burden reducers," see Laudan (2016, p. 110-37).

[38] In some ways, the project is similar to objective probabilistic accounts of legal proof in that the conditions for its implementation seem to be impossible to realize in practice. See Pardo (2023a, p. 444-46).

[39] Note that there is no analytical reason why error allocation must be located in one rule (the standard of proof) as opposed to two, three, ten, or a hundred rules. See Pardo (2007, p. 372). Laudan assumed this reduction because, in his view, it would be easier as a practical matter to calculate and implement. See Laudan (2015). But, for the reasons discussed above and in the next section, the reduction

3.2. The illusion of precision and the explanatory structure of proof

Laudan's analysis assumed a kind of precision for standards of proof that (1) appears illusory at best and (2) ignores the explanatory structure of proof.[40] Laudan critiqued the BARD standard for its imprecision and its vague language (as well as its non-epistemic focus).[41] In its place, Laudan (2016, p. 18-45) never settled on a specific verbal formulation, but he argued that the standard should be informed by the various frequencies and harms associated with both false convictions and false acquittals (including potential increases in violent crime). His analysis on this issue assumed that a precise criminal standard of proof could be constructed that would produce the desired ratio of errors (Laudan, 2016, p.80). Accordingly, he then used this analysis to argue that the BARD standard in the United States is too high and that the standard of proof in criminal cases should be lowered.[42]

Laudan's analysis on this issue is inconsistent with features of legal proof discussed in Part 2. In particular, the analysis is inconsistent with the judge-jury relationship in the United States and with the role of standards of proof in evaluating sufficiency of evidence. As discussed in Part 2, the context for sufficiency-of-evidence review involves a minimal standard of epistemic permissibility, within an adversarial system.

One might argue that these inconsistencies are not a problem. They are not a problem, so the response might go, because Laudan was attempting to *change* law's epistemic practices not *explain* them. Not exactly, in my view. What the potential response misses is that Laudan was attempting to change specific rules and aspects of legal proof—while leaving basic structural features of legal proof in place, including the jury trial and the judge-jury relationship.[43] Here is the central problem with this approach: a standard of proof can either be implemented with precision or not—which for Laudan meant implemented based on the *epistemic features of the evidence*, not subjective beliefs or credences—to produce a desired ratio of errors (for example,

---

seems to reimpose similar kinds of impossible conditions that it was meant to avoid. Moreover, even with the same standard of proof, the number and types of errors will vary from jurisdiction to jurisdiction. See Allen & Elliott-Smith (forthcoming).

[40] On the latter, see the discussion in Part 2.

[41] See note 28, *supra*.

[42] Laudan (2016, p. 101-105) argued that the BARD standard was responsible for too many false acquittals and generally made it too difficult to convict defendants who had committed violent crimes (who would then go on to commit additional violent crimes). Other scholars have critiqued Laudan on this issue. Gardiner (2017) critiques Laudan's use of empirical information on crime statistics. Dei Vecchi (2020) points to an internal tension between Laudan's claim that the BARD standard is vague, arbitrary, subjective, on one hand, and his claim that BARD is the cause of false acquittals based on a rational assessment of the strength of the evidence, on the other hand.

[43] Laudan was not analyzing legal epistemology in the abstract. The primary focus of his work in legal epistemology was on the United States legal system, with a particular focus on the proof process in criminal cases.

eight false acquittals for every false conviction, or whatever).[44] If a judge can determine with precision whether the evidence meets the standard of proof (for example, an objective probability threshold), then there is really nothing left for a jury to do. The judge should just apply the standard and produce the desired result.[45] But if, on the other hand, a judge cannot determine with precision whether a standard of proof is satisfied in a given case (which seems more plausible), then we are back to something like the current proof structure for evaluating the sufficiency of evidence: a minimal threshold in a context of epistemic permissibility (in which either outcome is permitted) that is dependent on the adversarial system (see Part 2). But if this is so, then it is unclear how the legal system could ever produce a precise ratio of errors, even with a precise standard of proof.

Relatedly, Laudan's analysis also failed to appreciate the *explanatory* structure of the proof process (outlined in Part 2). Laudan was familiar with what is termed "inference to the best explanation" (IBE) in the philosophy of science, its basic features, as well as its critiques and limitations in that context (Laudan, 1981). Laudan appeared to assume—mistakenly in my view—that the inferential reasoning process would operate similarly in both contexts (law and science). Therefore, he critiqued the possibility that inference to the best explanation could shed light on legal proof in criminal cases (Laudan, 2007). This is because, for example, something might be the best explanation but not proven BARD, or not the best explanation but enough to raise a reasonable doubt.

What Laudan's critique misses, however, is that the law can use and adapt similar ideas and concepts in ways that do not necessarily match their features in philosophy or science. In the United States, the explanatory thresholds for standards of proof, both during trial and in sufficiency-of-evidence review, do not necessarily match the explanatory thresholds in other contexts (Pardo & Allen, 2008; Allen & Pardo 2019a; Pardo 2023a). In particular, for higher standards of proof such as BARD, the explanations must be stronger than being the best one, or better than the alternatives. Correspondingly, a defense explanation may raise a reasonable doubt even if not the "best" one, so long as it is good enough to be plausible. This example illustrates how features of a concept or idea that might present challenges or worries in other philosophical contexts (such as the philosophy of science) may not be problems when applied to the law.[46] The example also illustrates how philosophical concepts and ideas might provide useful tools for illuminating or making explicit epistemic aspects of the law—even when they do not contain all the necessary features, or match the

[44] See Laudan (2006, p. 53) ("The principal question is not whether the jurors, individually and collectively, are convinced by the prosecution. The issue is whether the evidence they have seen and heard *should be* convincing in terms of the level of support it offers the prosecution's hypothesis").

[45] The only possible reason to have a jury trial might be to preserve the jury's power in the United States to nullify or acquit despite the evidence satisfying the standard of proof.

[46] For a discussion of how objections to IBE in the philosophy of science do not apply to law, see Pardo & Allen (2008, p. 243-45).

same uses, as in other philosophical contexts. To know whether they prove to be a useful tool one has to look closely at the details of particular legal systems.

From the discussion in this Part, three specific lessons emerge—as well a fourth, meta-lesson—for how epistemology may shed light on legal proof or how it may fail to do so. One lesson is the danger in trying to *reduce* or *isolate* features of legal proof that cannot be reduced or isolated (such as allocating the risk of error) from the system as a whole. The second lesson is the *illusion of precision*: the illusion is to think that some complex feature, rule, or practice (such as a standard of proof) can be given sharp boundaries when it cannot.[47] The third lesson is that a potential *mismatch* between philosophical concepts and how they arise in law is not necessarily a problem for law.[48] The fourth, which is a meta-lesson, is the risk of *mismodeling* or *misdescribing* the legal phenomena that are the object of epistemological inquiry.

The next Part discusses how these lessons apply to recent work in legal epistemology. Again, the discussion is from the perspective of an evidence scholar, exploring both (1) where legal epistemology illuminates or helps us understand something about legal proof, and (2) where, because of the lessons discussed above, epistemological discussions may go awry.

4. CONTEMPORARY LEGAL EPISTEMOLOGY

The lessons that emerged in the discussion of Laudan's scholarship also apply, in different ways and to various degrees, to current issues in the literature on legal epistemology. This Part discusses five examples: (1) statistical evidence, (2) the "preponderance of the evidence" standard, (3) epistemic safety, (4) knowledge, and (5) epistemic injustice.

4.1. Statistical evidence

The first example concerns vast (and rapidly growing) literature by philosophers concerning statistical evidence. Much of this literature is framed around solving or explaining the so-called "puzzle" or "paradox" that arises in well-known hypothetical cases or vignettes involving statistical evidence as well as explaining how, if at all, statistical evidence differs from non-statistical evidence.[49] Perhaps the most famous is the Blue Bus hypothetical in which the Blue Bus Company owns 70 percent of the

[47] The illusion of precision also applies to probabilistic accounts of standards of proof. See Pardo (2023a, p. 444-49).

[48] In other words, it is not a problem for law simply because the explanatory reasoning process in legal cases does not match the IBE process in epistemology or in the philosophy of science.

[49] One sign of the growth in the field of legal epistemology is that the number of recent articles on this topic is now too large to list them all. For overviews of the examples and issues, see Ross (2020); Gardiner (2019a); Pardo (2019).

buses in town, and there is no other evidence in a civil case other than a common assumption that a bus was responsible for an accident.[50] The "puzzle" or "paradox" is supposed to arise from the fact that the statistic (70 percent) appears to surpass the standard of proof (where the "preponderance of the evidence" standard is assumed to mean greater than 50 percent). Yet the intuition of most people is that the evidence is not *sufficient* to support liability (Wells, 1992, p. 740). The issue is typically further complicated with the assumption that an eyewitness testifying that a Blue Bus is responsible for the accident *is* sufficient to support liability, even when it is stipulated that the witness is 70 percent likely to be accurate. Much ink has been spilled trying to explain or explain away these hypotheticals.[51]

As philosophically interesting as this literature is, it does not, in my view, shed much light on legal proof (at least not in the United States). One reason, as Ron Allen and Christopher Smiciklas (2022) have shown, is that there is no general distinction between statistical and non-statistical evidence in United States case law. Therefore, the objects being explained or analyzed in the literature—which are typically intuitions about the hypothetical vignettes—do not actually map onto legal reality.[52]

---

[50] Other prominent examples include Gatecrashers (Cohen, 1977, p. 75) and Prisoners in the Yard (Nesson, 1979, p. 1192-93).

[51] I confess to being guilty of this as well. See, e.g., Allen & Pardo (2007); Pardo & Allen (2008); Pardo (2018); Pardo (2019).

[52] A clear example of this is Enoch & Spectre (2019) in response to Pardo (2018). Pardo (2018) provided a United States Supreme Court case as a counter-example to the authors' previous claims about sensitivity and statistical evidence. In that case (Tyson Foods, Inc. v. Bouaphakeo, 136 S. Ct. 1036 (2016)), the Court (1) rejected any general distinction between statistical and individualized evidence, and (2) concluded that the plaintiffs' statistical-average evidence was sufficient to establish a liability claim for individual plaintiffs in the reference class. In response, Enoch & Spectre (2019, p. 183-84) explain that their account was not meant to apply to "statistical evidence" generally, but rather only to the "underlying intuitive concerns" in cases such as "Blue Bus and the phenomenon it is an example of. This is the phenomenon sometimes called base-rate evidence, sometimes market-share evidence, or sometimes naked statistical evidence." (p. 183-84). They attempt to distinguish *Tyson Foods* as a case "where statistics were admitted in order to prove something about the circumstances," whereas examples like Blue Bus involve "direct base-rate data for guilt or liability" (p. 183). This is an odd response for two reasons. First, qualifying the limited scope of their sensitivity account (i.e., that they are only explaining intuitions about hypotheticals such as Blue Bus and not actual features of the law) seems to be inconsistent with the less qualified, more general, issue that they purported to be addressing. See Enoch, Spectre & Fisher (2012, p. 199) ("The distinction between statistical and individual evidence is a general one, and it seems to call for a general solution."). Second, and more importantly, *Tyson Foods* was not a case in which statistics were used merely "to prove something about the circumstances." Rather, the case does seem to involve similar "underlying intuitive concerns" as in the hypothetical cases—in that a group statistical average was used to prove the central disputed factual issue in the case by applying that average to individuals within the reference class in order to establish civil liability. How this situation differs from hypotheticals such as Blue Bus is unclear (as Enoch and Spectre do not define the "statistical evidence" they are addressing beyond the above descriptions). But the *Tyson Foods* discussion does make clear that the target of their analysis is intuitions about the unrealistic hypothetical vignettes and not features of legal proof. See also Allen (2022) ("if real courts are not behaving consistently with the intuitions … it is a complete mystery what the point of the intuitions

But even taking the hypotheticals at face value, there is no general "puzzle" or "problem" in need of explanation if my starting picture of legal proof (see Part 2) is correct. Moreover, that picture of legal proof helps us to diagnose *why* there is no puzzle or paradox. There is thought to be a puzzle or paradox because of the combination of two premises (1) that the ownership statistic (70 percent) is the "probative value" of the evidence, and (2) that the standard of proof is a probabilistic threshold (50 percent) (Pardo, 2019, p. 255). But if neither premise is true, then there is no problem or paradox. And neither premise is true in the picture presented. In that picture, the standard of proof imposes an explanatory threshold and depends on the competing explanations, and probative value is determined by how well the evidence supports or challenges the explanations. Instead of "puzzles" or "paradoxes," the hypotheticals are merely unspecified cases, without enough information to decide one way or the other. This might explain why, despite all the ink that has been spilled, there is no consensus on a particular solution to the cases.

The literature in legal epistemology on statistical evidence exhibits several of the limitations discussed in Part 3 and, thus, the lessons that emerged in that discussion also apply to much of the literature on statistical evidence. First, it is overly reductive in trying capture something essential about statistical evidence (and in assuming that the category can be meaningfully separated from other kinds of evidence). Second, it relies on an illusion of precision (i.e., that probative value and the standard of proof can be sharpy defined). Third, it often misses the explanatory structure of proof (and thus mismodels details regarding probative value and standards of proof). The discussions may be valuable philosophically for several reasons internal to epistemology,[53] but they do not illuminate much about evidence law and legal proof.

Putting aside the specific issue of statistical evidence, the literature on legal epistemology has generated a number of other interesting and potentially illuminating ideas about standards of proof and features of legal evidence. I turn to some of these in the examples to follow.

4.2. The "preponderance of the evidence" standard

The second example concerns philosophical accounts of the "preponderance of the evidence" (POE) standard. The literature on legal epistemology has devoted con-

---

might be … both market-share and base-rate evidence are potentially admissible evidence … in every jurisdiction in the United States; thus, precisely what is being distinguished from what and why is somewhat opaque."). The idea that there is something called "naked" statistical evidence is an unfortunate distraction in the literature. There is no such thing, and the idea has caused additional confusion. See Pardo (2019, p. 262 n.128)

[53] For example, similarities and differences between the legal hypotheticals and the "lottery" hypotheticals in epistemology may reveal additional issues that are relevant to philosophy. See also Buchak (2014); Moss (2018).

siderable attention to the topic of legal standards of proof; however, much of this literature mismodels the object of inquiry by missing several of the features outlined above in Part 2. These features include (1) the explanatory structure of the proof process; (2) the manner and context in which courts assess the sufficiency of evidence; and (3) the role of the parties in supplying contrasting explanations.

Two notable exceptions in the epistemological literature—by philosophers Sarah Moss (2023) and Georgi Gardiner (2019b), respectively—illuminate, in different ways, aspects of legal proof. They also come closer than several other philosophical accounts to accurately describing the POE standard within the context of legal proof.[54] Each account adapts a "relevant alternatives" framework from epistemology—which connects with the fact that legal proof is *comparative* in the sense that it depends on comparing alternative possible explanations of the disputed events.[55] Both accounts, however, also fail to accurately capture some of the important underlying legal details.

The preponderance standard in theory is meant to treat the parties roughly equally with regard to the risk of error.[56] It purports to do this by favoring whichever side the evidence better supports.[57] In practice, legal fact-finders choose the better of the competing explanations advanced by the parties (Pardo & Allen, 2008; Allen & Pardo, 2019a). Fact-finders may (but are not obligated to) consider alternatives not advanced by the parties. And courts review the sufficiency of evidence by evaluating whether a reasonable jury could find a party's explanation better than the alternative(s) that are or were advanced (Pardo 2023a).

According to Moss's (2023) account of legal proof: a fact is proven under the preponderance standard when the fact-finder has *knowledge* that the disputed fact has a probability greater than 0.5 (p. 199). In other words, the jury must *know* that "the defendant is *probably* liable" (Moss, 2023.).[58] This account diverges from the pre-

---

[54] In the philosophical literature, the POE standard is typically assumed to mean a greater than 0.5 probability (without necessarily specifying the type of probability at issue). For problems with this assumption, see Pardo (2023a); Allen & Pardo (2019a). As discussed above, this assumption helps to create the appearance of a "paradox" for statistical evidence. Importantly, jury instructions on the POE standard vary considerably among jurisdictions, and they generally do not use "0.5" or "50 percent" language. Leubsdorf (2015) surveys different instructions, some of which are consistent with, and some of which are inconsistent with, "0.5" or "50 percent" language.

[55] Along with fitting the explanatory structure of legal proof (discussed in Part 2), the "relevant alternatives" framework provides additional epistemic support for explanationist accounts of legal proof, to the extent the former is itself epistemically justified. Moss's account also usefully illustrates possible connections between knowledge and legal proof. For a philosophical account of the BARD standard that fits with aspects of the explanatory structure of legal proof, see Lackey (2021).

[56] See E.M.D. Sales, Inc. v. Carrera, 604 U.S. 45, 50 (2025) ("the preponderance-of-the-evidence standard has remained the default standard of proof in American civil litigation. That default makes sense: The preponderance standard allows both parties in the mine-run civil case to share the risk of error in roughly equal fashion.").

[57] See Pardo (2023a).

[58] See Moss (2023, p. 199) ("Here is the thesis I want to defend: proof of liability by a preponderance of the evidence requires that the factfinder know that the defendant is *probably* liable"). See also

ponderance standard, however, both practically and theoretically. As a practical matter, the account does not match what fact-finders or courts do on a regular basis. The law may consider a fact proven by a preponderance of the evidence, and the evidence may be held to be sufficient, in the absence of such knowledge.[59] And, theoretically, such an account fails to treat the parties equally with regard to risk of error.[60] Moss's account, in other words, places too much risk on the party with the burden of proof: a fact-finder could find a party's explanation is better (and thus more likely), and a judge could think this finding is reasonable—which, as discussed in Part 2, is all that is necessary—and yet the party would still lose under the Moss account whenever the fact-finder does not *know* that the probability is greater than 0.5.[61]

Gardiner's (2019b) account does not require probabilistic knowledge or quantify the preponderance standard. Instead, it requires that the party with the burden produce sufficient evidence "to rule out preponderant error possibilities" (p. 298). The alternative explanations that must be ruled out are the ones that are (1) "most significant" or "most predominant"; (2) the "most usual ways claims such as the one being made are false;" (3) "the ones a judge would most expect to see;" and (4) those "a virtuous inquirer would think most important and that seem most weighty as potential doubts" (p. 301).

Similar to Moss's account (2023), Gardiner's account requires too much of parties with the burden of proof, and, consequently, places too much of the risk of error

---

ibid. at 200 ("the defendant is proven liable by this standard only if the judge or jury has greater than .5 credence that he is

liable. The knowledge account adds an objective element to this condition: the defendant is proven liable by a preponderance of the evidence if and only if this probabilistic belief constitutes knowledge."). One conceptual advantage to Moss's account of the preponderance standard is that it preserves a role for *knowledge* while allowing for the possibility for false positive verdicts. In other words, knowing that a defendant is *probably liable* is consistent with the defendant not, in fact, being liable. See note 3, *supra*. Building a knowledge requirement into the BARD standard raises additional conceptual issues, as Moss discusses (see p. 204-09).

[59] For examples, see Pardo (2023a). The preponderance standard also applies to *admissibility-of-evidence* decisions in addition to the elements of civil causes of action. See *supra* note 38. Courts routinely apply the *Federal Rules of Evidence* (for example, finding that a hearsay exception is satisfied or that a prior act can be admitted for a non-character purpose) in the absence of such probabilistic knowledge.

[60] See note 60, *supra*.

[61] Under this account, the allocation of the risk of error would be a problem for more than civil plaintiffs. The POE standard also applies to some affirmative defenses that criminal defendants must prove. Thus, criminal defendants would fail under this account despite offering a more persuasive case than the prosecution on the issue. And the preponderance standard governs disputes about the *admissibility* of evidence for all parties in both civil and criminal cases. See notes 38 and 63, *supra*. In most instances, judges and juries do not have knowledge of the probability of the disputed fact, conditioned on the evidence. For reasons discussed below (see Part 4.4), I share with Moss's account an affinity for *knowledge* as potentially illuminating aspects of legal proof. See also Pardo (2018); Pardo (2010); Pardo (2005). However, in my view, knowledge (or something like knowledge) helps to illuminate a goal or aim of legal proof—as opposed to *knowledge being a requirement of standards of proof* (as it appears to be under Moss's account).

on that party. It does so because it requires that for a fact to be proven by a preponderance of the evidence, the fact-finder is obligated to "rule out" alternative possible explanations beyond those advanced by the parties. First, in United States, under the "party presentation" and "master of the case" principles, courts and fact-finders are not required to consider alternative explanations and arguments not advanced by the parties.[62] And second, when there is a choice to be made between plausible alternatives advanced by the parties, there is no necessary requirement that the alternative be "ruled out" (only that one explanation is found, comparatively, to be more plausible).[63]

There may be normative reasons why the law ought to adopt one of these two proposed epistemic standards, but they each fail as a *descriptive* or *explanatory* account of the preponderance standard (at least as it is deployed in the United States). The accounts capture some, but not all, of the explanatory structure of legal proof.[64] And they miss details about the adversarial principles that structure legal proof in the United States. They also each allocate the risk of error in a manner that deviates from the standard's underlying policy goal.

4.3. Epistemic safety

The third example involves the use of epistemic safety (and the related but distinct notion of sensitivity) as a tool to explain features of legal evidence. Much of this literature has focused on debates about statistical evidence. For the reasons discussed in Part 4.1, the analysis below puts the specific issues regarding statistical evidence to the side and focuses on a more general issue. To what extent can safety or sensitivity help to explain micro-level features of legal evidence such as probative value?[65] In answering this question, two of the lessons discussed in Part 3 will be important to keep in mind. First, a concept or idea need not fully match its features and uses in

---

[62] Fact-finders are often *permitted* to consider alternative explanations not advanced by the parties (unless, for example, a legal rule or instruction precludes consideration of an alternative). But fact-finders are not required or obligated to do so. See note 17, *supra*. See also Clark v. Sweeney, 607 U.S. __ (2025) ("In our adversarial system, we follow the principle of party presentation... To put it plainly, courts call 'balls and strikes'; they don't get a turn at bat.") (internal citations omitted).

[63] See Part 2. Indeed, rather than elucidating the POE standard, a need to "rule out" alternatives (or do more than provide the better explanation) implicitly raises the standard of proof from POE to a higher standard of proof such as "clear and convincing evidence." For a detailed illustration of how this works in the context of a recent United States Supreme Court opinion on racial gerrymandering, see Pardo (forthcoming).

[64] In this respect, they come closer than other philosophical accounts in illuminating legal standards of proof. See also Lackey (2021).

[65] See note 4, *supra*. A third influential modal account, see Smith (2018), based on normalcy conditions (or the set of possible worlds that fact-finders would normally expect to see) is generally outside the scope of the discussion. I note, however, that the normalcy account potentially mismodels legal proof (based on the picture presented in Part 2) to the extent it is inconsistent with (1) the minimal threshold for evaluating sufficiency of evidence, and (2) reliance on the parties to construct the set of relevant contrasting explanations. See also note 20, *supra*.

philosophy in order to potentially illuminate legal proof. Second, a philosophical concept or idea may help to explain an aspect of legal proof even when that aspect (e.g., a standard of proof or probative value) cannot be fully reduced or explained in terms of that concept or idea.

In epistemology, safety and sensitivity are typically used to analyze *true* beliefs—starting with a true belief and asking whether the belief would continue to be held in various counterfactual scenarios (or possible worlds). Sensitivity concerns the closest possible worlds in which something is false, and explores whether an epistemic agent would still believe it to be true. Safety concerns how easily the belief could be false.

As with the discussion of Laudan and "inference to the best explanation," applications of safety and sensitivity to legal proof will not match the uses and features of these concepts in other philosophical contexts. In legal settings, we typically do not know whether a legal verdict or some item of evidence (e.g., a witness identification) is true or false. But we can ask, *assuming it were true*, would some evidence, or an inference from that evidence, be safe or sensitive?[66]

A second important difference between legal proof and other uses in epistemology concerns how the relevant possibilities are determined and measured. In philosophy, the possible worlds that count as "close" or "closest" can be tricky and contested. But in legal proof, the most important possibilities at issue are the alternative explanations advanced by the parties, even if they would not necessarily count as the "closest" worlds in philosophical analyses.

With these two differences in mind, I have argued that something like epistemic safety—even if it is not isomorphic with safety in epistemology—affects the probative value of evidence (Pardo, 2018). Consider an example. Suppose that we have an eyewitness to a crime who identifies a suspect's photo as the perpetrator. But suppose further that the eyewitness would identify any photo placed before him as the perpetrator. The identification has low probative value as evidence (even if the identification was correct in this instance).[67] In part, this is because the evidence and

[66] Of course, one is free to object that what is being discussed is not really "safety" or not really "sensitivity," just as (in Part 3) one is free to object that the process is not really "inference to the best explanation." From the perspective of understanding law (although perhaps not from the perspective of philosophy), the response should be: *so what?* What matters, from the perspective of law, is whether the idea being discussed reveals something about an aspect of legal proof, regardless of whether it diverges from its features or uses elsewhere. Although, for the sake of clarity, it will often be preferable to qualify uses of the terms (e.g., "something like safety") and point out relevant differences.

[67] Epistemic safety is distinct from the "likelihood ratio" (a common measure that purports to capture aspects of probative value). The likelihood ratio compares the probability of the evidence given one hypothesis (e.g., the defendant's guilt) with the probability of the evidence given the negation of that hypothesis (e.g., the defendant's innocence). For differences between probability and epistemic safety, see Pritchard (2015). One important difference is that some possibilities with low probabilities of occurring may nevertheless be unsafe (i.e., could easily be true). For limitations with the "likelihood ratio" as a general theory or definition of probative value, see Allen & Pardo (2023); Allen & Pardo (2007). This is not to suggest that "likelihood ratios" or other probabilistic measures do not also illuminate aspects of

inferences from it are unsafe—the evidence and the inferences a fact-finder draws from it could have so easily been false. As the philosopher John Greco (2012, p. 193) explains: epistemic "safety *just is* reliability throughout a space of close counterfactual situations". And something like "reliability throughout a space of close counterfactual situations" is a potentially important consideration when assessing the strength or quality of legal evidence in distinguishing between contrasting explanations. [68] This is an especially important consideration in the context of legal proof when the parties are arguing about which possibility is the true one.

By contrast, epistemic sensitivity does not have much of an effect on the probative value of legal evidence. This is because the "closest" possible world in which the evidence, or an inference from the evidence, is false, *may not be a close world* (Pardo 2018, p. 65). When this is the case, the evidence may have high probative value even though it is insensitive. To continue with the eyewitness example, suppose instead that the only circumstance in which the eyewitness would have made a false identification is if there were an elaborate disguise to make someone else look like the true perpetrator (otherwise, the witness would not have made a false identification). Even if the witness *would* make a false identification in that far-fetched scenario, this would not lower the probative value of the evidence (unless perhaps this is the contrasting explanation advanced by the opposing party, with evidence to support). And, indeed, one reason why the evidence retains its high probative value is because the evidence, and inferences from it, are safe. The evidence and inferences could not easily be false. [69]

---

probative value. Both safety and probability affect probative value (along with other contextual factors such as a party's need for the evidence; the evidentiary alternatives; and the competing explanations of the parties; as well as the general reliability of the evidence).

[68] In a critique of Pardo (2018), Enoch and Spectre (2019, p. 187) argue that invoking this property (possessed by safety but not sensitivity) "begs the question" absent an argument that this is "an important desideratum, legally and epistemically." But, as explained in Pardo (2018, p. 71-74), this is an important desideratum in the context of legal proof precisely because it affects the probative value of evidence. In the context of legal proof—which is comparative and focused on a small subset of contrasting explanations (see Part 2)—reliability in close possibilities is *one thing* that affects probative value. Enoch and Spectre's critique misses this connection: "Pardo often accuses Sensitivity of being a poor guide to *probative value*, but it is unclear what he means by that phrase (p. 188)." "Probative value" was not used idiosyncratically—it was used in accord with its uses throughout evidence law in the United States, as was explained throughout. See, e.g., Pardo (2018, p. 53, 57, 71-72). They continue: "Sometimes it seems that he just takes high probative value to be the property of good evidence. If so, accusing Sensitivity of having poor probative value is of course a serious accusation, but it needs substantiating" (p. 188). As explained in the text above (and in Pardo, 2018, p. 75), sensitivity often has little effect on probative value *precisely because* it lacks the property that safety possesses (reliability in close possible scenarios). In misunderstanding the relationships between safety, sensitivity, and probative value, their critique does not address the thesis of Pardo (2018) that safety has a greater effect than sensitivity on the probative value of legal evidence (and thus plays a great role throughout evidence law). To be fair, Enoch and Spectre were focused on a different theoretical project (i.e., explaining intuitions about hypothetical vignettes involving some subset of "statistical evidence"). See note 52, *supra*.

[69] See Pardo (2018, p. 75) ("Sensitivity does not track the reliability of evidence, or the risk of drawing erroneous inferences, in close possible worlds. Safety, however, does track these features and

In this example, something like safety helps to illuminate an aspect of probative value. And it does so even though the application does not match some of the features and uses of safety in epistemology. In this sense, the use of safety in the context of legal proof is similar to the discussion of "inference to the best explanation" in Part 3. As in that discussion, a concept or idea may diverge from its other philosophical uses while fitting with a feature of legal proof. Moreover, safety might illuminate an aspect of probative value *even though there is more to probative value than safety*.[70] In other words, the concept might illuminate one aspect of legal proof without capturing all there is to probative value. To assume otherwise is to adopt too reductive of an approach[71]—similar to Laudan and standards of proof—when applying epistemology to legal proof.

---

thus has a greater effect on the probative value of evidence."). Epistemic sensitivity (or something like it), however, directs our attention to other potentially important features of evidence (such as whether the evidence would exist in the close possible scenarios in which the disputed fact was false). Sensitivity may, for that reason, produce some incentive effects (see Enoch, Fisher & Spectre, 2012). But for a critique of the incentive-effect arguments for sensitivity, see Allen (2022, p. 268-71).

[70] See Pardo (2018, p. 71) ("Safe evidence may have low probative value for other reasons. For example, where there are not close possibilities in which the findings are false, then evidence of otherwise poor quality will be safe. In addition, safe evidence may have low probative value because it is cumulative of other evidence"). Moreover, other modal properties (for example, "normic support" in Smith (2018)) may also affect probative value.

[71] For an example of such a reductive approach, see Johnson King (2022). Johnson King critiques Pardo (2018) but the critique misinterprets the arguments in Pardo (2018) in two important ways. The critique mistakenly assumes that safety was put forward as both (1) a *general* account of probative value, and (2) a *necessary* requirement for admissibility. Neither assumption is true. Based on these mistaken assumptions, Johnson King argues that examples of (1) problematic but safe evidence, and (2) admissible but unsafe evidence, provide counter-examples to the arguments in Pardo (2018). Because of the mistaken assumptions, neither point is correct. On the first point, see the quote in note 70, *supra*. Thus, examples of safe but otherwise problematic legal evidence are not counter-examples to the thesis (and indeed were already acknowledged). On the second point, see Pardo (2018, p. 53) ("To be clear, I do not claim that safety provides a complete account of legal evidence or all aspects of the law of evidence. Rather, I argue for the more modest claim that safety is an important consideration for legal evidence and that it plays a greater explanatory role than sensitivity.") and *ibid*. (p. 72) ("This is not to suggest that safety provides an overarching theory of admissibility.") Thus, pointing to examples of admissible evidence that are unsafe likewise does not challenge the thesis. Johnson King's reductive approach also seems to assume that applications of philosophical concepts to law must match their uses in other philosophical contexts in order to illuminate law. See Johnson King (2022, p. 10-14) (arguing that the factivity of safety and sensitivity makes them inappropriate for explaining legal evidence). See note 66, *supra*. Johnson King's analysis also mismodels the legal structure for sufficiency of evidence (what the analysis terms "adequacy" and "inadequacy," see p. 18-21); this structure, as explained above, is (1) permissive, (2) comparative, and (3) can be met with statistical evidence. See Part 2 and, e.g., note 52, *supra*.

4.4. Knowledge

The fourth example concerns the concept of knowledge itself. Similar to epistemic safety and IBE, aspects of legal proof may share some features with knowledge (as it is understood in epistemology) while lacking other features of the concept. Despite this mismatch, knowledge may help to illuminate aspects of legal proof.

The traditional conception of knowledge as (1) justified, (2) true, (3) belief has analogs when applied to legal verdicts (Pardo, 2010, p. 38-42). First, legal verdicts have truth or factual accuracy as a central goal or aim. Second, legal verdicts require some level of epistemic justification (Ho, 2008, 89-99). But this level need not necessarily be the same level as required for knowledge (Redmayne, 2008).[72] Third, legal verdicts involve some level of epistemic commitment by the fact-finder. But this commitment need not necessarily be a belief—it may instead involve some type of acceptance or acceptability threshold (see Ferrer Beltran, 2006).

In epistemology, it is generally accepted that these three conditions may not be *sufficient* for knowledge. As famously shown by Edmund Gettier (1963), they will not be sufficient if there is an accidental (or coincidental) gap between the justifying evidence, on one hand, and the truth of the belief, on the other hand. In these situations, someone arrives at a true conclusion by accident or coincidence, despite having what appears to be justifying evidence for that belief. I have argued that a similar gap or accidental connection may also undermine legal verdicts (Pardo 2010; Pardo 2011). For example, consider a criminal defendant, framed by the police with falsified evidence (for example, planted cocaine), who is then convicted based on this falsified evidence (for example, for possession of cocaine). But suppose also that this defendant happens to be guilty of that crime (i.e., the defendant did happen to knowingly possess undiscovered cocaine at the time of the arrest). The verdict is true, and there was sufficient evidence to support or justify it at trial, but the verdict is still problematic or defective.[73] Drawing out the similarities between knowledge

[72] For discussions, see Ho (2008, p. 89-99); Redmayne (2008).

[73] The example assumes that the falsified nature of the evidence was not apparent at the time of trial. In a recent article, Ho (2025, p. 12-13) challenges the conclusion that the verdict in the example was true ("*Pace* Pardo, the verdict is false in *Framed Defendant*. The defendant was not in illegal possession of the drugs that were allegedly found in his car. He was unaware of their existence until they were revealed by the police officers."). According to Ho, the fact-finder must find the defendant "guilty of the particularized instance of the type of the offence cited in the charge and as alleged by the prosecution." He suggests an analogy: "Take an even more far-fetched scenario: I am charged with the murder of A. I didn't do it. The police set me up. If I am convicted, the verdict is false. We do not consider the verdict true because, 'by luck' and unbeknownst to the judge, I happened to have murdered B. My conviction is for the murder of A and not B." This critique, however, misinterprets the nature of legal verdicts (in the United States), and the analogy does not work. The propositions that the fact-finder (jury or judge) are asked to find BARD are the *elements* of the crime. In drug-possession cases, the element refers to a *type of substance* (not, as Ho argues, the "particularized instance" of it). See, e.g., United States Courts

and legal proof makes explicit this aspect of legal proof that might otherwise go unnoticed. Successful legal verdicts require not only sufficient evidence and truth; they also require an appropriate connection between the evidence and truth.[74]

In drawing attention to this largely unnoticed issue, epistemology helps to shed light on legal proof. Notice, this insight applies even if successful verdicts do not constitute knowledge for other reasons (for example, having to do with justification or belief).[75] Thus, knowledge provides another example—as with inference to the best explanation—in which a philosophical concept may illuminate an aspect of legal proof, even when it does not match its features or uses elsewhere in philosophy.

4.5. Epistemic Injustice

A growing body of recent scholarship has applied philosophical work on epistemic injustice to legal proof (Tuerkheimer, 2017; Epstein & Goodman, 2019; Gonzales

---

for the Ninth Circuit, Manual of Modern Criminal Jury Instructions (2010) (listing the first element as "First, the defendant knowingly possessed [*specify controlled substance*]."); see also *ibid*., Comment ("The defendant does not need to know what the controlled substance is so long as the defendant knows that he or she has possession of such a substance... See also United States v. Soto-Zuniga, 837 F.3d 992, 1004-05 (9th Cir. 2016) (knowledge of type and quantity of drugs not element of offense)"). Therefore, in the example, the jury will be asked to the decide whether "the defendant knowingly possessed cocaine." This proposition, *pace* Ho, is true in the example (i.e., the defendant did knowingly possess cocaine at the time of arrest). Nor does Ho's analogy ("murder of A and not B") work as a *reductio* or a counter-example because murder is person specific. See *ibid*. (listing the first element for "murder—second degree" as "First, the defendant unlawfully killed [*name of victim*]"). Thus, in Ho's example, this proposition *is* false.

[74] For example, in cases of newly discovered evidence or in cases with defective trial evidence, it is not enough for the prosecution to point to other (not admitted) evidence showing that the verdict was true (or likely to be true BARD). Rather, the justifying evidence must have been internal to the proof process at trial. Ho (2025, p. 13) suggests that in such circumstances (i.e., accidentally true outcomes) the law should "breathe a sigh of relief... Luckily, all turned out well." Ho argues that if newly discovered evidence shows the original verdict was both not justified and "we think the verdict may well be false," then the defendant may be entitled to some "legal recourse … perhaps even a retrial." But the law (in the United States) goes further than this. On appeal, if the evidence is insufficient to support the conviction, then the question of whether the verdict is thought to be true or false is irrelevant—the verdict is constitutionally defective and will be overturned. See Jackson v. Virginia, 443 U.S. 307 (1979). If newly discovered evidence later shows that what appeared to be sufficient trial evidence was defective, then the defendant is entitled to a retrial, even if other evidence external to the trial shows the original verdict was true. See Pennsylvania v. Reese, 663 A.2d 206 (1995).

[75] See Pardo (2010, p. 54-55) (discussing other ways in which successful verdicts may fail to qualify as knowledge). See also note 61, *supra*. Nor does the above analysis depend on any proposed "solution" to the Gettier problem. Moreover, although the analysis is framed in terms of the traditional JTB conception of knowledge, the analysis also applies to "knowledge first" approaches in epistemology (in which knowledge is taken to be a basic or fundamental mental state used to explain other epistemic concepts such as justification, rather than knowledge itself being composed of other more basic epistemic properties), see Williamson (2000). For other possible connections between knowledge and legal proof, see Moss (2023); Mueller (2024).

Rose, 2021; Picinali, 2024). This work provides another useful example of the ways in which ideas from epistemology can illuminate overlooked and underexplored aspects of legal evidence and proof.[76] In this final example, I discuss some of the positive ways in which a focus on epistemic injustice contributes to understanding legal proof and illustrate how some of the lessons discussed in the previous sections apply in this context.

Much of this scholarship builds on the pioneering work of Miranda Fricker (2007) and focuses on the specific category of epistemic injustice known as "testimonial" injustice.[77] Within the literature, central examples of testimonial injustice involve situations in which a witness is (1) assigned a level of credibility that is lower than is warranted by the evidence (i.e., a "credibility deficit"), and (2) this credibility deficit is due to identity-based prejudices, stereotypes, or assumptions (for example, based on race or gender).[78] The "primary harm" of epistemic injustice is that the agent is "wronged in one's capacity as a knower … a capacity essential to human value" (p. 44). Fricker (2007) refers to this primary harm as an "intrinsic injustice,"[79] but also discusses other practical consequences and epistemic harms that may also follow from this primary harm.[80] Scholars have also expanded on this category by illustrating how, in the context of legal evidence and proof, credibility "excesses" may also lead to cases of epistemic injustice (Lackey, 2023; Gonzales Rose, 2024). The feature of legal proof that makes the latter a genuine concern—perhaps more so than in other social contexts—is that legal proof is *comparative* (see Part 2) and of-

[76] Epistemic injustice also provides a potential counter-example to the claim by Enoch, Fisher & Spectre (2021, p. 92-3) that, as a normative matter, the law should not sacrifice *any* amount of accuracy to secure an epistemic property. In other words, some readers may have the intuition that it may be worth sacrificing a very *small* amount of overall accuracy if it means removing *large* amounts of epistemic injustices from a legal system. The issue is likely a contested one, in which the underlying details will matter significantly.

[77] A distinct but related type of epistemic injustice is "hermeneutical injustice," which involves "the injustice of having some significant area of one's social experience obscured from collective understanding owing to hermeneutical marginalization." Fricker (2007, p. 158). In the context of legal proof, the two types of epistemic injustices may interact in complex ways (for example, hermeneutical injustices may interact with legal evidence to produce testimonial injustices). The discussion in this section focuses on relatively simple cases of testimonial injustice.

[78] Fricker distinguishes "identity based" epistemic injustices from "innocent errors" in assessing credibility evidence (p. 21). See also *ibid*. ("The conception of identity in question concerns the most important part of an individual's social identity (racial, political, sexual, religious), where the affiliation tends, in our culture at least, to be experienced by the individual as essential to who she or he really is."). For a discussion of types of epistemic injustice beyond those that arise from identity prejudices, see Lackey (2023).

[79] ("The form that this injustice takes specifically in cases of testimonial injustice is that the subject is wronged in her capacity as a giver of knowledge. The capacity to give knowledge to others is one side of the many-sided capacity so significant in human beings: namely, the capacity for reason. We are long familiar with the idea … that our rationality is what lends humanity its distinctive value. No wonder, then, that being insulted, undermined, or otherwise wronged in one's capacity as a giver of knowledge is something that can cut deep." Fricker (2007, p. 44).

[80] See Fricker (2007, p. 46-48).

ten zero-sum.[81] In other words, *overvaluing* evidence in one direction will affect the *opponent* of the evidence (similar to the ways in which *undervaluing* evidence harms the *proponent* of the evidence).

Two of the lessons discussed in the preceding sections apply to epistemic injustice in the context of legal evidence and proof. First, in ways that parallel the discussion of knowledge and legal proof (see Part 4.4), the category of testimonial injustice is, in fact, broader than scholars have typically assumed. Second, in ways that parallel the discussion of Laudan and standards of proof (see Part 3), the relationship between instances of epistemic injustice and the accuracy of case outcome is sometimes more uncertain and complex than is typically assumed. Each point is discussed in turn.

First, when is there testimonial injustice in the context of legal proof? The paradigm case is a credibility "deficit" because of an identity-based prejudice. Because of an identity-based prejudice, stereotype, or assumption a fact-finder ascribes a level of credibility to a witness that is lower than what is rationally warranted by the evidence. Scholars have expanded this category by illustrating how credibility "excesses" can also produce testimonial injustices (Lackey, 2023; Gonzales Rose, 2024).[82] Credibility deficits and excesses both assume that a fact-finder has ascribed a level of credibility to a witness that *deviates from the level that is rationally warranted by the evidence* (Fricker, 2007).[83] In other words, the very idea of a credibility *deficit* or *excess* depends, as a necessary feature, on the assumption that there is some "correct" level of credibility that is epistemically warranted or justified by the evidence.[84] Suppose that we have a witness to a crime and that the epistemically warranted level of credibility for this witness is level *X*.[85] Suppose a fact-finder ascribes a level of *X* credibility to the witness (and thus

---

[81] For a discussion of this aspect, see Gonzales Rose (2024).

[82] Lackey (2023) also expands the category beyond *identity-based* epistemic injustices.

[83] Fricker is clear that the norm at work is for the hearer (in the context of legal proof, the fact-finder) to match credibility to what is indicated by the evidence. Fricker (2007, p. 19) ("credibility is a concept that wears its proper distribution on its sleeve... the hearer's obligation is obvious: she must match the credibility she attributes to her interlocuter to the evidence that he is offering the truth.").

[84] This is an important, but sometimes overlooked, point. As Fricker has explained: credibility assessments are not just about social *power* (*i.e.*, who appears and does not appear to be credible)—they are also about what is *true* as a matter of fact, and what is *rationally warranted* based on the evidence. Indeed, Fricker (1998) argues that the very ideas of epistemic *discrimination* and epistemic *injustice* presuppose that there is a fact of the matter about credibility that the hearer or fact-finder gets wrong. (See ibid., p. 175) ("The possibility of identifying a given set of working-indicator-properties as *discriminatory* depends on being able to say that they manifest the norm of credibility so as to attribute rational authority where it should not, and/ or withhold it where it should not. If rational authority were the same as the power to appear rationally authoritative, *then there could be no genuine notion of discrimination*, and the political perspective in epistemology would have been lost before it was won." [emphasis added]). See also Picinali (2024, p. 26) ("Of course, if testimonial injustice occurs, the level of credibility assigned to the testimony is by definition inaccurate, in that it is lower than is warranted by the evidence.").

[85] See Gonzales Rose (2024, p. 181-82) (using quantified examples to illustrate how credibility shifts may occur during a trial).

accordingly relies or does not rely on the testimony).[86] But, suppose the fact-finder ascribed level *X* credibility based on identity-based stereotypes or prejudices and not based on a rational assessment of the evidence. In this example, there is no credibility *deficit* (or *excess*), but there is a testimonial injustice. The "primary harm" of testimonial injustice has occurred—the witness has been wronged in their capacity as a knower, for an identity-based reason.[87] The fact-finder ascribed the "correct" amount of credibility but did so for identity-based, discriminatory reasons. This is so even though the fact-finder landed on the "correct" amount coincidentally or by accident.[88] Thus, the category of testimonial injustice is broader than typically assumed in that it may occur even in the absence of a credibility deficit or excess.

This conclusion parallels the knowledge discussion in the previous section. There, we saw that, at the *macro-level*, a coincidental or accidental connection between the evidence and the truth could undermine a legal verdict. This might occur when the fact-finder reaches a verdict, based on misleading evidence, that coincidentally turns out to be true. Here, at the *micro-level*, a similar coincidental or accidental connection between credibility assessments and credibility evidence can result in an epistemic injustice.

Second, what is the relationship between instances of testimonial injustice and the accuracy of case outcomes? Again, the primary harm in testimonial injustice is to the witness *qua* knower.[89] Scholars have also emphasized, however, the significant practical consequences that may follow from testimonial injustice in the legal context (Tuerkheimer, 2017; Lackey, 2023; Gonzales Rose, 2024). For example, acts of testimonial injustice may cause jurors to incorrectly assess the probative value of testimony and thus draw mistaken inferences, leading potentially to fact-finding errors.[90] The most straightforward examples of testimonial injustice involve parties and victims—for example, criminal defendants whose testimony is not believed because of testimonial injustice or the victims of crime not believed because of testimonial injustice. In ways that parallel the discussion of Laudan and standards of proof (see Part 3), however, the consequences that instances of testimonial injustice have on the accuracy of case outcomes often cannot be isolated or separated from the sys-

---

[86] Note that even properly assessing credibility evidence may lead to fact-finding errors. A witness may be speaking truthfully and accurately even if a rationally warranted assessment of the credibility evidence indicates that they should not be believed, and vice versa.

[87] See notes 78, 79, and 80, *supra*, and accompanying text.

[88] Picinali (2024, p. 26) discusses how testimonial injustices that involve credibility deficits may lead to conclusions that are "serendipitously accurate" because they are part of "overall assessments" that are corrected by other evidence in the case. My point in the text is broader—there can be a testimonial injustice without any credibility deficit (or excess) to begin with.

[89] See notes 79 and 80, *supra*, and accompanying text.

[90] Note that fact-finding errors may occur even when fact-finders properly assess credibility evidence. See note 90 *supra*. Moreover, the credibility evidence that is available, even if properly assessed, may not be very probative of the witness's credibility. And fact-finders may fail to properly assess credibility evidence for reasons other than testimonial injustice ("innocent errors," see Fricker, 2007, p. 21).

tem as a whole.[91] This means that the practical consequences of epistemic injustice in particular cases, or in kinds of cases, can present complex and often uncertain empirical questions.[92] In some cases, genuine acts of testimonial injustice may correct or cancel out other injustices.[93] For an extreme example, if we assume an innocent criminal defendant on trial, then testimonial injustice toward a prosecution witness (a credibility "deficit") or testimonial injustice in the form of a credibility "excess" given to a defense expert witness, may, in fact, push the jury toward a factually correct result.[94] As with standards of proof, philosophical analysis and applications can illuminate or draw our attention to overlooked possibilities and draw out analytical implications. But it is too reductive to assume—as with standards of proof taken in isolation—that practical effects on the accuracy of case outcomes will follow from the analysis.[95] How issues of epistemic injustice arise in particular cases and in particular legal systems—and thus how the legal system should respond—requires a detailed understanding of how they interact with the system as whole.

5. CONCLUSION

This article presented a complicated picture of the relationship between legal epistemology and legal proof. The picture was presented from the perspective of an evidence scholar, asking how legal epistemology might help us better understand legal evidence and proof. From this perspective, a number of general lessons emerged.

---

[91] See, *e.g.*, Tuerkheimer (2017) (explaining how epistemic injustice interacts systemically with crime rates, crime-reporting rates, false-allegation rates, and police and prosecution behavior to affect the cases that go trial).

[92] This assumes that one has a case of testimonial injustice in the first place. For some of the empirical difficulties in establishing individual instances of testimonial injustice, see Arcila-Valenzuela & Paez (2024). Discussions of testimonial injustice typically illustrate the phenomena by assuming the acts of injustice affect the parties or victims in a straightforward manner. See, e.g., Gonzales Rose (2024, p. 180) ("Let us suppose the defendant is Black and the police officer is White"). Testimonial injustice, however, may also occur in more complex relationships between parties and witnesses. For example, in Pena-Rodriguez v. Colorado, 137 S. Ct. 855 (2017), the United States Supreme Court held that the use of racial stereotypes or animus to convict a defendant may violate the defendant's constitutional right to a fair trial. The case largely focused on comments made by a juror about the defendant. However, a less-discussed aspect of the case is that the juror also made similar statements about the defendant's alibi witness (including falsely stating the witness was "an illegal" and thus not credible). This act of testimonial injustice would have existed even if the juror had not made statements about the defendant, and even if the defendant had been a different race. See also Roberts (2023, p. 645-46).

[93] They may also compound. See, *e.g.*, Lackey (2023) (explaining how the same criminal defendant may be given a credibility excess and a credibility deficit at different stages of the criminal process).

[94] What is true in this one case may also occur systematically, depending on the base rates of true and false claims; the quality of the evidence (both credibility and non-credibility evidence); and whether the fact-finder properly assesses the admissible evidence; among other factors. The point is that the relationship between testimonial injustice and accuracy depends on complex empirical issues that cannot be isolated from the system as whole. See also Allen & Elliott-Smith (forthcoming).

[95] See Part 3.

One lesson is the danger in trying to reduce or isolate features of legal proof that cannot be reduced or isolated. We saw this lesson in the discussion of Laudan on standards of proof and the risk of error, as well as in the discussions of statistical evidence, safety, and epistemic injustice. A second lesson is the illusion of precision—which arises when thinking that a legal feature, rule, or practice can be given sharp boundaries when it cannot. We saw this lesson in the discussions of standards of proof and probative value throughout the article, and in particular in the discussion of statistical evidence. The third lesson is that a potential mismatch between how philosophical concepts are defined and used in philosophy, on one hand, and how they arise in law, on the other hand, may not necessarily be a problem for law. We saw this lesson in the discussions of inference to the best explanation, safety, sensitivity, and knowledge. The analysis also drew attention to a fourth, meta-lesson: the risks of mismodeling or misdescribing the legal phenomena that are the objects of epistemological inquiry. We saw this in the discussions of standards of proof throughout the article—including Laudan on the BARD standard and more recent attempts by philosophers to explain the POE standard—as well as the discussions of statistical evidence, safety, and knowledge.

The vision of legal epistemology presented in this article is not one of simply applying or deducing consequences or implications from a favored philosophical theory or concept.[96] Rather, it requires a close attention to the epistemic details of

[96] I thus share with Enoch, Fisher, & Spectre (2021, p. 85) an objection to doing legal epistemology "just by doing epistemology" ("What we focus on in this paper—and what we say is deeply mistaken—is the attempt, roughly, to do legal epistemology just by doing epistemology, or to treat epistemological considerations (say, about the nature of knowledge or evidence) as decisive, or even just intrinsically relevant, to normative questions in evidence law theory."). The vision of legal epistemology presented in this article is outside the scope of their critique in that epistemology was used to explain or illuminate aspects of legal proof. Moreover, some of the applications in this article are broader than—and do fit comfortably inside—the "two-tiered" model they provide for legal epistemology. They argue that epistemology might be relevant either for "designing the evidence law regime" or for a "conscientious fact-finder" (p. 85). But a central example used throughout this article, sufficiency-of-evidence review by trial and appellate courts, does not seem to fit either category—thus suggesting a broader possible scope for legal epistemology than they suggest. Furthermore, the vision presented in this article shares with Larry Laudan (see note 23, *supra*) the view that epistemic considerations are ultimately relevant for how they may contribute (or hinder) the law's goals regarding factual accuracy and the minimization of errors. Thus, I reject some of the framing in Enoch, Fisher, & Spectre (2021, p. 89-92) to the extent it posits tradeoffs between truth and some epistemic considerations. They present a thought experiment: a normative choice between (1) a more accurate legal system, or (2) a less accurate system that secures an additional epistemic criterion or property (e.g., knowledge)—and they argue that it is a kind of "epistemic fetishism" to prefer the latter. But we can run their thought experiment with the concept of "evidence" itself (because being "supported by evidence" is itself an epistemic consideration). So, if we follow their framing: should we prefer a legal system with more accuracy or one with less accuracy that requires that verdicts be supported by evidence? This strikes me as a strange question. See also Roush (2023, p. 124) ("an externalist about knowledge will find the presented choice puzzling because for us it is inconceivable that knowledge would not itself include insurance of accuracy"). Although I share with Enoch, Fisher, and Spectre the assumption that any epistemic considerations in the legal system that *do* result in decreased overall accuracy need to be justified in light of some other legal, political, or moral value (such as fairness, equality, or justice) and not in light of epistemology alone.

the evidentiary proof process of a particular legal system (or legal systems). In this sense, in applying epistemology to law, the vision presented resembles something like the philosophy of physics (Wallace, 2021; see, e.g., Norton, Pooley & Read, 2023). The similarities involve the combination of abstract philosophical tools and concepts, on one hand, with detailed specialized knowledge of complex phenomena, on the other. The complexities that arise from this interaction are also an invitation for more collaboration between philosophers and legal scholars about the details from each discipline that matter. I conclude by returning to Larry Laudan. Laudan, in my view, seemed to endorse a similar approach. Rather than deduce consequences for law from theories in the philosophy of science, he grappled with the messy details of legal proof, criminal law, and criminal procedure in the United States and attempted to bring some clarity to what he observed.

BIBLIOGRAPHY

Allen, R. & Elliott-Smith, N. (forthcoming). Proof Beyond a Reasonable Doubt Doesn't Exist: Except as an Emergent Property of a Complex System. *Journal of Criminal Law & Criminology*.

Allen, R., Pardo, M., Lawrence, W. & Smiciklas, C. (2025). Minimal Rationality and the Law of Evidence. *Journal of Criminal Law & Criminology*, *115(2)*, 269-315.

Allen, R. & Pardo, M. (2023). Evidence, Probability, and Relative Plausibility: A Response to Aitken, Taroni & Bozza. *The International Journal of Evidence & Proof, 27(2)*, 126-42.

Allen, R. (2022). Naturalized Epistemology and the Law of Evidence: A reply to Pardo, Spellman, Muffato, and Enoch. *Quaestio Facti, 3,* 253-72.

Allen, R. & Smiciklas, C. (2022). The Law's Aversion to Maked Statistics and Other Mistakes. *Legal Theory, 28(3),* 179-209.

Allen, R. & Pardo, M. (2019a). Relative Plausibility and Its Critics. *The International Journal of Evidence & Proof, 23 (1-2)*, 5-59.

Allen, R. & Pardo, M. (2019b). Clarifying Relative Plausibility: A Rejoinder. *The International Journal of Evidence & Proof*, *23(1-2)*, 205-17.

Allen, R. & Laudan, L. (2008). Deadly Dilemmas. *Texas Tech Law Review, 41(1),* 65-92.

Bentham, J. (1843). *An Introductory View of the Rationale of Judicial Evidence, Works VI(5)*, (Bowring ed., 2002).

Allen, R. & Pardo, M. (2007). The Problematic Value of Mathematical Models of Evidence. *The Journal of Legal Studies, 36(1),* 107-40.

Amaya, A. (2019). The Explanationist Revolution in Evidence Law. *The International Journal of Evidence & Proof*, *23(1-2)*, 60-67.

Arcila-Valenzuela, M. & Paez, A. (2024). Testimonial Injustice: The Facts of the Matter. *The Review of Philosophy and Psychology, 15*, 582-602)

Buchak, L. (2014). Belief, Credence, and Norms. *Philosophical Studies, 169(2)*, 285-311.

Cheng, E. & Pardo, M. (2015). Accuracy, Optimality, and the Preponderance Standard. *Law, Probability, and Risk*, *14(3)*, 193-212.

Cohen, L. (1977). *The Probable and The Provable*. Oxford University Press.

Dei Vecchi, D. (2020). Laudan's Error: Reasonable Doubt and Acquittals of Guilty People. *The International Journal of Evidence & Proof, 24(3)*, 211-32.

Enoch, D., Fisher, T. & Spectre, L. (2021). Does Legal Epistemology Rest on a Mistake? On Fetishism, Two-Tiered System Design, and Conscientious Fact-Finding. *Philosophical Issues, 31(1),* 85-103.

Enoch, D. & Spectre, L. (2019). Sensitivity, Safety, and the Law: A Reply to Pardo. *Legal Theory, 25(3),* 178-199.

Enoch, D., Spectre, L. & Fisher, T. (2012). Statistical Evidence, Sensitivity, and the Legal Value of Knowledge. *Philosophy & Public Affairs, 40(3)*, 197-224.
Epstein, D. & Goodman, L. (2019). Discounting Women: Doubting Domestic Violence Survivors' Credibility and Dismissing Their Experiences. *University of Pennsylvania Law Review*, *167(2),* 399-461.
Ferrer Beltran, J. (2006). Legal Proof and Fact Finders' Beliefs. *Legal Theory, 12(4),* 293-314.
Fricker, M. (2007). *Epistemic Injustice*. Oxford University Press.
Fricker, M. (1998). Rational Authority and Social Power: Towards a Truly Social Epistemology. *Proceedings of the Aristotelian Society, New Series, 98,* 159-77.
Ho, H. (2025). Should we be Convicting People We Don't Believe to be Guilty? *Quaestio Facti, 6,* 101-29.
Ho, H. (2008). *A Philosophy of Evidence Law*. Oxford University Press.
Gardiner, G. (2019a). Legal Burdens of Proof and Statistical Evidence. D. Coady & J. Chase (eds.) *The Routledge Handbook of Applied Epistemology*. Routledge.
Gardiner, G. (2019b). The Reasonable and the Relevant. *Philosophy & Public Affairs, 47(3)*, 288-318.
Gardiner, G. (2017). In Defence of Reasonable Doubt. *Journal of Applied Epistemology, 34(2)*, 221-41.
Gettier, E. (1963). Is Justified True Belief Knowledge? *Analysis, 23(6),* 121-23.
Gonzales Rose, J. (2024). A Critical Perspective on Testimonial Injustice: Interrogating Witnesses' Credibility Excess in Criminal Trials: A Comment on Frederico Picinali's "Evidential Reasoning, Testimonial Injustice and the Fairness of the Criminal Trial." *Quaestio Facti, 7,* 173-85.
Gonzales Rose, J. (2021). Race, Evidence, and Epistemic Injustice. C. Dahlman, A. Stein & G. Tuzet (eds.). *Philosophical Foundations of Evidence Law*. Oxford University Press.
Greco, J. (2012). Better Safe than Sensitive. K. Becker & T. Black (eds.) *The Sensitivity Principle in Epistemology*, 193-206.
Johnson King, Z. (2022). Safety, Sensitivity, and Admissibility. *Synthese, 200(6)*, 1-22.
Lacey, N. (2004). *A Life of H.L.A. Hart: The Nightmare and the Noble Dream*. Oxford University Press.
Lackey, J. (2023). *Criminal Testimonial Injustice*. Oxford University Press.
Lackey, J. (2021). Norms of Criminal Conviction. *Philosophical Issues, 31(1),* 188-209.
Laudan, L. (2006). *Truth, Error, and Criminal Law: An Essay in Legal Epistemology.* Cambridge University Press.
Laudan, L. (2016). *The Law's Flaws: Rethinking Trials and Errors?.* College Publications.
Laudan, L. (2015). Why Asymmetric Rules of Procedure Make it Impossible to Calculate a Rationally Warranted Standard of Proof. Available at: *https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2584658* .
Laudan, L. (2007). Strange Bedfellows: Inference to the Best Explanation and the Criminal Standard of Proof. *The International Journal of Evidence & Proof, 11(4)*, 292-306.
Laudan, L. A Confutation of Convergent Realism. *Philosophy of Science, 48(1)*, 19-49.
Leubsdorf, J. (2015). The Surprising History of the Preponderance Standard of Civil Proof. *Florida Law Review, 67(5)*, 1569-1619.
Mikhail, J. (2007). "Plucking the Mask of Mystery From its Face": Jurisprudence and H.L.A. Hart. *Georgetown Law Journal, 95*, 733-779.
Moss, S. (2023). Knowledge and Legal Proof. T. Gendler *et al.* (eds.). *Oxford Studies in Epistemology, Volume 7*. Oxford University Press.
Moss, S. (2018). *Probabilistic Knowledge*. Oxford University Press.
Mueller, A. (2024). Legal Proof: Why Knowledge Matters and Knowing Does Not. *Asian Journal of Philosophy, 3(28)*, 1-22.
Nesson, C. (1979). Reasonable Doubt and Permissive Inferences: The Value of Complexity. *Harvard Law Review, 92(6)*, 1187-1225.
Norton, J., Pooley, O. & Read, J. (2023). The Hole Argument. *Stanford Encyclopedia of Philosophy*, available at: *https://plato.stanford.edu/entries/spacetime-holearg/* .
Pardo, M. (forthcoming). The Standard of Proof in Racial-Gerrymandering Cases: *Alexander's* Mistake and its Implications. *University of Illinois Law Review, 2026(1).*

Pardo, M. (2023a). What Makes Evidence Sufficient?. *Arizona Law Review, 65(2)*, 431-78.
Pardo, M. (2023b). On Proving Mabrus and Zorgs. *Vanderbilt Law Review, 76(6)*, 1653-80.
Pardo, M. (2019). The Paradoxes of Legal Proof: A Critical Guide. *Boston University Law Review, 99(1)*, 233-90.
Pardo, M. (2018). Safety vs Sensitivity: Possible Worlds and the Law of Evidence. *Legal Theory, 24(1),* 50-75.
Pardo, M. (2013). The Nature and Purpose of Evidence Theory. *Vanderbilt Law Review*, *66(2)*, 547-613.
Pardo, M. (2011). More on the Gettier Problem and Legal Proof: Unsafe Nonknowledge Does Not Mean that Knowledge Must be Safe. *Legal Theory, 17(1)*, 75-80.
Pardo, M. (2010). The Gettier Problem and Legal Proof. *Legal Theory, 16(1),* 37-57.
Pardo, M. & Allen, R. (2008). Juridical Proof and the Best Explanation. *Law and Philosophy, 27*, 223-68.
Pardo, M. (2007). On Misshapen Stones and Criminal Law's Epistemology. *Texas Law Review, 86(2)*, 347-83.
Pardo, M. (2005). The Field of Evidence and the Field of Knowledge. *Law and Philosophy, 24(4)*, 321-92.
Picinali, F. (2024). Evidential Reasoning, Testimonial Injustice and the Fairness of the Criminal Trial. *Quaestio Facti, 6,* 201-35.
Pritchard, D. (2015). Risk. *Metaphilosophy, 46(3),* 436-61.
Psillos, S. (2023). In Science We Trust: Larry Laudan (1941-2022), *Journal for General Philosophy of Science, 54*, 523-33.
Quine, W. (1969). *Ontological Relativity & Other Essays*. Columbia University Press.
Redmayne, Mike. (2008). Exploring the Proof Paradoxes. *Legal Theory, 14(4),* 281-309.
Roberts, P. (2023). Theorising Evidence Law. *Oxford Journal of Legal Studies, 43(3),* 629-49.
Ross, L. (2020). Recent Work on the Proof Paradox. *Philosophy Compass, 15(6)*, 1-11.
Roush, S. (2023). Knowledge, Evidence, and Naked Statistics. L. Oliveira (ed.). *Externalism About Knowledge*. Oxford University Press.
Smith, M. (2018). When Does Evidence Suffice for Conviction?. *Mind, 127(508)*, 1193-1218.
Tuerkheimer, D. (2017). Incredible Women: Sexual Violence and the Credibility Discount. *University of Pennsylvania Law Review, 166(1),* 1-58,
Twining, W. (2019). Bentham's Theory of Evidence: Setting a Context. *Journal of Bentham Studies, 18(1)*, 20-37.
Wallace, D. (2021). *Philosophy of Physics: A Very Short Introduction*. Oxford University Press.
Wells, G. (1992). Naked Statistical Evidence of Liability: Is Subjective Probability Enough? *Journal of Personality & Social Psychology, 62(5)*, 739-52.
Williamson, T. (2000). *Knowledge and its Limits*. Oxford University Press.

Case law and legislation

Addington v. Texas, 441 U.S. 418 (1979)
Bourjaily v. United States 483 U.S. 171 (1987)
Clark v. Sweeney, 607 U.S. __ (2025)
Colorado v. New Mexico, 467 U.S. 310 (1984)
Cruzan v. Missouri Dep't of Health, 497 U.S. 261 (1990)
E.M.D. Sales, Inc. v. Carrera, 604 U.S. 45 (2025)
Federal Rules of Evidence
Federal Rules of Civil Procedure
Federal Rules of Criminal Procedure
Greenlaw v. United States, 554 U.S. 237 (2008)
Huddleston v. United States, 485 U.S. 681 (1988)

In re Winship, 397 U.S. 358 (1970)
Jackson v. Virgina, 443 U.S. 307 (1979)
Pena-Rodriguez v. Colorado, 137 S. Ct. 855 (2017)
Pennsylvania v. Reese, 663 A.2d 206 (1995)
Reeves v. Sanderson, 530 U.S. 133 (2000)
Tyson Foods, Inc. v. Bouaphakeo, 136 S. Ct. 1036 (2016)
United States v. Sineneng-Smith, 140 S. Ct. 1575 (2020)
United States v. Soto-Zuniga, 837 F.3d 992, 1004-05 (9th Cir. 2016)
United States v. Fatico, 458 F.3d 388 (E.D.N.Y. 1978)
United States Courts for the Ninth Circuit, Manual of Modern Criminal Jury Instructions (2010), available at: https://www.ce9.uscourts.gov/jury-instructions/

ENSAYOS

Quaestio facti. Revista Internacional sobre Razonamiento Probatorio
Quaestio facti. International Journal on Evidential Legal Reasoning
Sección: Homenaje a Larry Laudan
2026 I 10 pp. 223-243
Madrid, 2026
DOI: 10.33115/udg_bib/qf.i10.23162
Marcial Pons Ediciones Jurídicas y Sociales
© Sebastián Sandoval Ayala
ISSN: 2604-6202
Recibido: 31/05/2025 | Aceptado: 08/12/2025 | Publicado *online*: 12/01/2026
Editado bajo licencia Reconocimiento 4.0 Internacional de Creative Commons

LA PRUEBA DE LOS HECHOS FUTUROS. CONSIDERACIONES SEMÁNTICAS Y EPISTÉMICAS PARA LA TOMA DE DECISIONES*

Sebastián Sandoval Ayala**

RESUMEN: toda disciplina debe lidiar con problemas de vocabulario que surgen cuando las categorías disponibles, los términos que utilizamos en nuestro lenguaje cotidiano, no resultan satisfactorias para llevar a cabo una empresa investigativa con la precisión que se estima necesaria. Producto de este desajuste discursivo, las comunidades disciplinarias debaten si las categorías sospechosas deben ser eliminadas o pueden ser redefinidas en términos menos problemáticos. En este trabajo, se estudia la categoría de hecho futuro. Luego de explicar por qué la categoría resulta sospechosa, se sostiene que ella no puede ser eliminada o redefinida en los términos propuestos por los teóricos de la prueba sin perder parte importante de su significado. Esto es relevante porque el significado perdido resulta indispensable para controlar las decisiones judiciales que se basan en consideraciones prospectivas, como ocurre al determinar el monto de un daño futuro o al decretar una medida cautelar como la prisión preventiva.

PALABRAS CLAVE: derecho probatorio; prueba; hecho; error; verdad.

---

* Una versión preliminar de este trabajo fue presentada en las Jornadas Nacionales de Derecho Probatorio celebradas en la Pontificia Universidad Católica de Valparaíso durante el año 2024. Agradezco especialmente el apoyo y los comentarios de los profesores Rodrigo Coloma y Jorge Larroucau en las distintas etapas de este trabajo.

** Doctorando en Derecho, Universidad Alberto Hurtado. Magíster en Derecho, mención en Derecho Público, de la Universidad de Chile. Profesor del Departamento de Teoría y Filosofía del Derecho de la Universidad Diego Portales. https://orcid.org/0000-0002-9355-5907

THE PROOF OF FUTURE FACTS. SEMANTIC AND EPISTEMIC CONSIDERATIONS FOR DECISION MAKING

ABSTRACT: Every new discipline must deal with some vocabulary problems that arise when the available categories, the terms we use in our everyday language, are not satisfactory to carry out a research enterprise with the precision deemed necessary. As a result of this discursive mismatch, disciplinary communities debate whether suspect categories should be eliminated or can be redefined in less problematic terms. In this paper, the category of future facts is studied. After explaining that the expression can be understood as a suspect category in evidence law, it is argued that it cannot be eliminated or redefined in the terms proposed by evidence theorists without losing an important part of its meaning. This is relevant because the lost meaning is indispensable to control judicial decisions that are based on prospective considerations, such as when determining the amount of a future damage or decreeing a precautionary measure such as preventive detention.

KEYWORDS: evidence law; proof; fact; error; truth.

SUMARIO: 1. CUESTIONES PRELIMINARES.— 2. UNA CATEGORÍA SOSPECHOSA.— 3. ¿CONOCIMIENTO SOBRE EL FUTURO?: 3.1. Altas expectativas epistémicas.— 4. EL PROBLEMA DE HABLAR SOBRE EL FUTURO: 4.1. Futuro incierto y futuro abierto; 4.2. Futuro y error.— 5. CONCLUSIONES Y ALGUNOS DESAFÍOS PENDIENTES.— BIBLIOGRAFÍA

1. CUESTIONES PRELIMINARES

En el vocabulario de los juristas podemos distinguir diversas categorías de *hechos* (Hunter, 2015). Se habla de hechos lejanos en el tiempo, internos, convenidos por las partes, públicos y notorios, entre otros. Dentro de este elenco de categorías, una de las que parece suscitar problemas especiales es la de *hecho futuro* (Muñoz, 2001, p. 85). Pese a que la expresión ha sido empleada de manera regular por la doctrina, existen dudas sobre su sentido o alcance[1]. Las sospechas parecen indicar que los términos que componen la expresión (hecho y futuro) denotan cierto tipo de contradicción que sería indicativa, en última instancia, de un mal uso del lenguaje.

Este supuesto mal uso del lenguaje presupone un uso correcto de la expresión que coincide, en el contexto probatorio, con llamar *hecho* exclusivamente a los eventos que han ocurrido o están ocurriendo en el mundo físico. Algunos teóricos de la prueba adscriben a este significado estricto cuando sostienen que los hechos ocurren fuera del proceso, que son independientes de las decisiones judiciales, que pueden ser captados sensorialmente y que constituyen la referencia semántica de

[1] Algunas de estas dudas pueden consultarse en: Rivera (2011, pp. 51-52); Mora (2021, p. 56 y ss.); Valenzuela (2022, pp. 172-173); y Bravo (2023, p. 87).

algunas de nuestras expresiones. El significado es estricto porque elimina algunos de los sentidos que el término conserva en otros contextos discursivos, reduciendo su utilidad, pero ganando precisión metodológica[2]. No caben dudas de que los hechos futuros no son hechos en este sentido de la expresión del mismo modo en que tampoco lo serían los *hechos negativos*, los *hechos convenidos por las partes* o los *hechos condicionales*.

Toda disciplina novel debe lidiar con algunos problemas de vocabulario que surgen cuando las categorías disponibles, los términos que utilizamos en nuestro lenguaje cotidiano, no resultan satisfactorias para llevar a cabo una empresa investigativa con la precisión que se estima necesaria[3]. Este proceso de ajuste discursivo requiere realizar varias operaciones. En ocasiones, los términos disponibles son redefinidos de tal manera que puedan cumplir con las expectativas de los investigadores. En otras, algunos términos deben ser eliminados porque contribuyen a generar o mantener confusiones mediante *preguntas tramposas*, que no pueden responderse sin incurrir en graves errores conceptuales. Los teóricos de la prueba suelen optar por la primera opción. Las redefiniciones de los «hechos definidos mediante términos valorativos» que ensayan Gascón (1999) y Taruffo (2002) constituyen un buen ejemplo. Se toma una categoría que no coincide con el programa investigativo de la teoría de la prueba y se la despoja de aquellos elementos diferenciadores que, prima facie, la caracterizan en el discurso de los juristas. Lo que se conserva es un mínimo común, el núcleo duro de lo que constituye el *hecho en sentido estricto*, que podría, ahora sí, ser considerado como un genuino objeto de prueba.

Respecto de los hechos futuros se podría pensar que tal núcleo duro lo componen aquellos hechos pasados o presentes que, proyectados en el tiempo, pueden generar consecuencias que poseen relevancia jurídica. Pero el objeto de prueba serían los hechos pasados o presentes que sirven de sustento a la proyección y no, en sentido estricto, los *hechos futuros* (Ferrer, 2007, pp. 32-35). La expresión deviene, entonces,

---

[2] La precisión metodológica se consigue identificando con claridad un objeto de estudio que permite uniformar la manera en que se produce conocimiento al interior de una disciplina, distinguiendo un set de preguntas o tópicos comunes, así como las herramientas o métodos de estudio que anticipan resultandos interpretables a partir del mismo aparato conceptual (Coloma, 2016, p. 273). En el marco de los estudios sobre la prueba, ello permite construir un límite entre los asuntos de prueba en sentido estricto y los asuntos de calificación normativa (Bayón, 2010, p. 9).

[3] Los desacuerdos sobre el uso de las categorías tienen su origen en la necesidad de recurrir al conocimiento común al momento de formular y transmitir una idea. De esta forma, toda categoría escapa parcialmente al dominio de quien la utiliza o, como señala Giannini (2021, p. 32), «se vive bajo la sospecha de que las palabras se nos escapan o nos llevan más allá de donde queremos ir o más allá del poder que hemos delegado en ellas. De allí el temor: sensación de distanciamiento, de inseguridad respecto de lo que significan o pueden llegar a significar por su cuenta». En su afán por escapar de este temor, las disciplinas científicas luchan por *domesticar* un conjunto de categorías. Esto implica eliminar y/o traducir las expresiones dudosas por otras que, prima facie, resultan menos problemáticas. Las categorías, por su parte, presentan cierto grado de resistencia cuando se afianzan en prácticas discursivas que se han solidificado en el uso del lenguaje cotidiano o en las prácticas discursivas que desarrollan los participantes de una disciplina. (Amsterdam y Brunner, 2002, pp. 27-37).

en un uso aproximativo del lenguaje. No probamos el futuro en sentido estricto, sino que hacemos "*como si*" para los efectos de justificar cierto tipo de decisión[4].

Pero redefinir conceptos es también una empresa riesgosa. Puede ocurrir que el sentido que se estima como más preciso en realidad encubra una notable pérdida de significado[5]. Esta pérdida de significado podría implicar que aquello de lo que antes estábamos conscientes (como, por ejemplo, de la dificultad de probar hechos futuros o la necesidad formular consideraciones prospectivas al tomar una decisión) ahora no esté presente dentro de la discusión fáctica. El peor escenario posible ocurre cuando esta pérdida de significado no puede ser capturada exitosamente por otras instancias de la discusión procesal, por ejemplo, porque tampoco corresponde a una genuina cuestión interpretativa o a una genuina cuestión de calificación normativa. Cuando esto ocurre, la precisión metodológica se consigue pagando un alto precio.

En este trabajo se analizan las sospechas que levanta la categoría de hecho futuro cuando se la analiza a partir de los conceptos propuestos por los teóricos de la prueba. Sostendré que parte importante de estas sospechas surgen de dos actitudes que vale la pena discutir. La primera es una actitud demasiado optimista respecto de aquello que podemos *saber* del pasado para los efectos de justificar una decisión en instancias judiciales. La segunda es una actitud demasiado cauta respecto de aquello que podemos *decir* sobre el futuro con relativa seguridad. Intentaré mostrar que ambas actitudes se vinculan directamente con la manera en que diferentes procesos judiciales afrontan el desafío de la incertidumbre, de tal manera que ajustar nuestras expectativas epistémicas (identificar aquello que podemos *saber*) constituye un paso necesario para mejorar la manera en que tomamos decisiones en instancias judiciales.

2. UNA CATEGORÍA SOSPECHOSA

La categoría de hecho futuro es una categoría sospechosa en tanto no coincide, de buena manera, con el aparato metodológico adoptado por parte importante de las disciplinas que, durante los últimos años, se han interesado por los problemas vinculados con la justificación externa de la premisa fáctica. Si bien se trata de una comunidad variopinta en que subsisten debates importantes, los miembros de estas

[4] La pregunta por el hecho futuro (es decir, por lo que podría llegar a ocurrir de adoptar ciertas decisiones en tiempo presente) es reemplazada por otras preguntas que se estiman menos problemáticas en tanto admiten una respuesta conforme al ideario de los participantes de la disciplina.

[5] Como señala Austin, «nuestro repertorio común de palabras encarna todas las distinciones que los hombres han creído conveniente trazar y todas las distinciones que han creído conveniente destacar durante la vida de muchas generaciones. No cabe duda de que es probable que tales distinciones y conexiones, puesto que han pasado el prolongado test de la supervivencia del más apto, sean más ricas, más sensatas y sutiles que las que cualquiera de nosotros podamos concebir en una tarde en nuestro sillón de trabajo». Esto, que es cierto respecto de las categorías del lenguaje cotidiano, probablemente tenga también asidero tratándose de las categorías de *hechos* que han desarrollado los juristas a lo largo del tiempo (Austin, 1971, p. 19).

disciplinas parecen agruparse en torno a un método de análisis común de corte racionalista, que gira en torno «a los problemas relacionados a la comprobación fáctica (la prueba) y la verdad» (Marchese, 2019, p. 11). El rótulo de *racionalista* se utiliza al menos en dos sentidos. En un sentido positivo, se usa para designar a quienes entienden que la actividad probatoria debe ser definida como una actividad de conocimiento «encaminada a averiguar la verdad sobre los hechos litigiosos, entendida esta bajo una noción correspondentista» (Accatino, 2019, p. 90). En un sentido negativo, se emplea para fijar un límite con otras concepciones de la prueba que adoptan o presuponen una noción no-correspondentista de la verdad; o que simplemente niegan la importancia de la verdad en los debates judiciales (Taruffo, 2013, p. 243).

Los teóricos racionalistas han manifestado una genuina preocupación sobre la manera en que se debe adquirir conocimiento en instancias judiciales para justificar racionalmente una decisión sobre los hechos probados. No pretendo discutir las principales tesis de esta concepción de la prueba, pero sí me interesa la noción de conocimiento que presumen cuando vinculan la actividad probatoria con la noción de verdad como correspondencia. A mi juicio, esta noción de conocimiento descansa en cierto tipo de empirismo que otorga un rol protagónico a la memoria como respaldo de la experiencia sensible o como fuente de las huellas sensoriales que un evento ha producido en el mundo físico[6]. De esta forma, las pretensiones que interesan a los teóricos racionalistas son aquellas «susceptibles de ser empíricamente verificadas o falsadas» (Tuzet, 2021, p. 85). Con ello se definen los deslindes de lo que constituye *prueba* en sentido estricto, una actividad que tiene como correlato y objeto a los hechos que han ocurrido en un mundo exterior al proceso judicial en que se discuten (Laudan. 2013, p. 36) y que operan como la referencia semántica de algunas de nuestras expresiones (Gascón, 2010, p. 49). En este sentido, Rivera (2011, p. 52) sostiene que los hechos futuros no pueden configurar un genuino objeto de prueba:

> En estos casos no se trata de probar, pues no hay registros o datos futuros que se trasladen del futuro al presente. Debe recordarse que la prueba es la realización de los medios probatorios que trasladan las fuentes en donde están estampadas las huellas de los hechos. Por ejemplo, nadie puede testimoniar que fue al futuro y vio a fulano de tal y conversó con él; o que la notaría del futuro registró un documento, o que en la clínica del futuro se trata al paciente tal. Eso es pura ficción.

[6] La sospecha relativa al conocimiento sobre los hechos futuros ha sido un tema recurrente en la filosofía de las ciencias naturales. Torretti (2014, p. 55), por ejemplo, identifica el origen de esta sospecha en la manera en que los empiristas suelen entender el conocimiento de los fenómenos contingentes: «como las impresiones futuras están aún por venir y las impresiones presentes se han convertido en pasadas antes de que uno siquiera tome nota de su presencia, desde un punto de vista humano el conocimiento descansa exclusivamente sobre la memoria. Pero un kantiano informado y alerta goza de una perspectiva diferente. Se da cuenta de que las impresiones sensoriales crudas solo nos son accesibles en los intersticios de una red conceptual, a la cual nos apresuramos a atarlas lo mejor que podemos (…) Ahora bien, los conceptos de objetos presentes suelen comprometer el futuro y nos dan un asidero sobre lo que está a punto de ocurrir». El guiño final de Torretti puede interpretarse como una invitación a considerar otras epistemologías posibles (diferentes al empirismo) que quizás nos provean de un mejor soporte conceptual para discutir sobre eventos futuros.

Nótese que las sospechas que levanta la categoría de *hecho futuro* son tanto semánticas como epistémicas. Las sospechas semánticas indican que no estamos en condiciones de llamar *hecho* a aquellos eventos que pueden no llegar a ocurrir en la realidad, para los efectos de fijar un genuino objeto de prueba en el marco de un proceso judicial[7]. Si vamos a discutir sobre la verdad o falsedad de un enunciado, los enunciados sometidos a escrutinio racional dentro del proceso deben ser de aquellos que puedan ser objeto de comprobación o falsación de acuerdo con los mecanismos propios de la epistemología general (Ferrer, 2021, pp. 22-23). De esta forma, Mora (2021, p. 68) ha destacado que «los hechos prospectivos, en sentido estricto, no son susceptibles de ser probados, en tanto no se puede corroborar algo que no ha sucedido».

Las sospechas epistémicas, en cambio, indican que no podemos adquirir conocimiento sobre eventos futuros para los efectos de justificar una decisión del modo y con la certeza requerida por el ordenamiento jurídico. En el centro del debate podemos identificar un conjunto de expectativas que se asocian a la partícula *hecho probado* para los efectos de justificar una decisión. Los enunciados que se postulan como hechos probados justifican la aceptación de una conclusión solo en la medida de que se validan conforme a las reglas que regulan la valoración probatoria. En los sistemas de sana crítica, esta validación es posible cuando una de la hipótesis se encuentra respaldada por inferencias que, en el mejor de los casos, conectan lo discursivo con lo sensorial, reduciendo la incertidumbre bajo el margen de lo tolerable por el sistema (Dei Vecchi, 2020, p. 28).

La categoría de *hecho futuro* es sospechosa en este sentido porque la manera estándar en que se produce conocimiento dentro de los procesos judiciales no sería la adecuada para justificar la adscripción normativa del estatus *probado* a un enunciado que trata sobre el futuro[8]. Este estatus se consigue mediante la *prueba,* una actividad que los cultores de la disciplina ya han definido en relación con la posibilidad de comprobar o falsar el contenido de un enunciado.

Pero queda todavía un asunto central al que no se ha prestado suficiente atención y que podría ser especialmente relevante para la discusión sobre el lugar que corresponde a los debates sobre el futuro dentro de la discusión probatoria (si es que corresponde alguno). Me refiero al concepto de error material[9]. Los teóricos de la prueba han sido enfáticos en señalar que su epistemología coloca en el centro del

[7] Siguiendo la definición propuesta por Russell (1968), Daniel González Lagier (2022, pp. 19-20) afirma que, en el marco de las discusiones sobre la prueba, entendemos que los *hechos* son aquellos eventos que vuelven verdaderas o falsas nuestras proposiciones o creencias.

[8] Uviller (1982, pp. 846-847), por ejemplo, sostiene que el discurso probatorio se estructura normalmente sobre ciertos verbos indicativos de que algo ha ocurrido y que ha dejado una huella en el mundo. Dentro de estos verbos podemos reconocer algunos tales como *detectar*, *inspeccionar*, *recolectar* y *resurgir*.

[9] El concepto de error material al que se refiere este apartado no es la única noción del error que manejan los teóricos de la prueba. Con todo, parece ser la más intuitiva. Al respecto, ver: Gascón (2005); Ferrer (2018); y Carbonell (2022).

debate la distinción entre lo probado y lo verdadero como un presupuesto necesario para sustentar el control de la actividad probatoria. Para ello, resulta relevante la noción de verdad como correspondencia. Si el objetivo de la actividad probatoria es determinar la verdad de lo ocurrido, entonces las decisiones judiciales que se sustentan en premisas falsas han de ser consideradas como «errores» [10]. El error surge de la discrepancia entre los hechos probados y los hechos verdaderos. En este sentido, Carbonell (2023, p. 77) sostiene que comprobar un error exige realizar un ejercicio de contraste entre lo que se declara en instancias judiciales y lo que ha ocurrido en una realidad ajena al proceso:

> desde un punto de vista epistémico, decir que un enunciado es erróneo equivale aproximadamente a afirmar que es falso. Ello significa que su calificación de "erróneo" implica una operación cognitiva de contraste entre el enunciado y un criterio. Asumiendo un concepto de verdad como correspondencia, este criterio consiste en la comprobación de la ocurrencia de hechos de los que da cuenta el enunciado en el mundo físico.

Pero ocurre que los enunciados mediante los que se expresan hechos futuros no pueden ser calificados como verdaderos o falsos al momento en que son proferidos. Esto porque tal calificación solo tiene sentido cuando los hechos que constituyen la referencia semántica del enunciado han, finalmente, ocurrido o no ocurrido dentro de los márgenes temporales que hemos definido para su comprobación o falsación. El problema de ello es que dicho momento de comprobación o falsación puede ocurrir solo una vez que el proceso judicial ha adquirido el estatus de cosa juzgada. Y si ello es así, entonces el error en el caso de los hechos futuros sería uno completamente irrelevante: no sería útil para modificar el resultado de un proceso. La categoría podría, entonces, ser sospechosa también en un sentido operativo. No serviría para sustentar el tipo de discusiones que nos interesan dentro del proceso judicial al momento de discutir sobre el error probatorio.

La noción de conocimiento presupuesta en el ideario de buena parte de los teóricos de la prueba establece, entonces, una conjunción entre las nociones de hecho probado, verdad como correspondencia y error material. El hilo conductor es empirista debido a que se asigna un rol protagónico a la memoria como sustento epistémico de la experiencia sensible al momento de representar la actividad probatoria. El fracaso de la actividad se condice con la falta de correspondencia entre lo probado, por un lado, y lo verdadero por otro. Prueba sin verdad sería una expresión que traspasa el margen de lo acordado por la disciplina en cuanto a su propuesta metodológica. En lo que sigue, intentaré mostrar que esta forma de entender el conocimiento no parece ser adecuada para los fines de controlar la calidad epistémica de las decisiones probatorias. Ello puede abrir paso a otras formas de entender el conocimiento producido en instancias judiciales que no necesariamente descansan en la noción de la verdad como correspondencia.

[10] Una discusión de esta tesis puede consultarse en Dei Vecchi (2023a).

3. ¿CONOCIMIENTO SOBRE EL FUTURO?

La distinción planteada entre hechos probados y hechos verdaderos resulta intuitiva a primera vista, pero constituye el punto de partida de la discusión que me interesa desarrollar para los efectos de precisar el sentido que podría tener una discusión sobre el futuro en instancias judiciales. Como podrá anticipar el lector, hablar de conocimiento sobre hechos futuros parece extraño si se acepta lo que hemos reseñado en el capítulo anterior. Esto ocurre porque los enunciados sobre hechos futuros no son comprobables, verificables o falsables en el sentido en que debiesen serlo para que la distinción entre lo probado y lo verdadero tenga sentido dentro de los márgenes temporales del proceso [11]. Si no contamos con hechos verdaderos al momento de tomar una decisión, no parece que tenga sentido hablar sobre la prueba como una genuina actividad de conocimiento dirigida a establecer la verdad sobre los hechos que son objeto de un litigio. La prueba de hechos futuros podría aceptarse solo en un sentido aproximativo o metafórico, pero no sería prueba en sentido estricto porque, del mismo modo, no versa sobre hechos en sentido estricto [12].

Ahora bien, de la afirmación según la cual *los enunciados sobre hechos futuros no pueden calificarse como verdaderos en sentido estricto* no se sigue necesariamente que todo ejercicio de prognosis sea equivalente a una adivinación injustificable en términos epistémicos. La tesis de que el conocimiento presupone verdad puede (y a mi parecer, debe) ser cuestionada al interior de la teoría de la prueba si esta pretende ser útil a la práctica judicial. Para sostener esta afirmación se pueden transitar dos caminos. El primero es discutir la utilidad de la distinción entre prueba y verdad para los efectos de representar la falibilidad de una decisión probatoria. Esto exige decir

[11] Es importante destacar que algunos enunciados sobre hechos futuros sí pueden ser verificados o falsados en el sentido estricto que presuponen los teóricos de la prueba cuando hablan sobre hechos pasados. El problema, como ya he anticipado, es que dicho ejercicio de verificación suele acontecer luego de que el proceso ha alcanzado el estatus de cosa juzgada material. Esto implica que, de verificarse una discrepancia entre lo afirmado en un primer momento y lo que podemos decir luego de ocurridos ciertos eventos, no sería una discrepancia relevante en términos procesales: no sería apta para modificar el estado o el resultado de la discusión jurídica que dio inicio al litigio.

[12] La dogmática ha utilizado la expresión hecho futuro para hacer referencia a enunciados sobre eventos que podrían no llegar a ocurrir. Los filósofos y epistemólogos que discuten sobre el futuro utilizan, de manera similar, la expresión "futuro contingente" para hacer referencia a instancias que, siendo posibles, no resultan necesarias (Øhrstrøm and Hasle, 2020). Lo que parece justificar el uso del término "hecho", en este contexto, es la posibilidad de sujetar la corrección de nuestras expresiones a lo que sabemos de la realidad, de tal forma que buscamos ajustar nuestro lenguaje al mundo, incluso si ese mundo todavía no ocurre (Briggs y Forbes, 2011, pp. 257-258). Los autores que discuten sobre futuros contingentes a menudo debaten si es posible afirmar verdades no triviales sobre el futuro (MacFarlane, 2003, p. 327) y si tal posibilidad nos compromete o no con un conjunto de tesis metafísicas sobre la naturaleza de instancias futuras (Torre, 2011, p. 364). En este trabajo, no desarrollaré esta discusión porque, a mi parecer, resulta discutible que la expectativa por encontrar la verdad sea adecuada para estructurar una discusión probatoria exitosa.

algo sobre las altas expectativas que los teóricos de la prueba asocian a la partícula de *hecho probado* en los debates judiciales. El segundo camino consiste en mostrar cómo diferentes ejercicios de prognosis pueden ser evaluados racionalmente incluso si no tenemos acceso a un *hecho verdadero* en sentido estricto. Esto exige decir algo sobre el tipo de fenómenos que nos interesa anticipar o proyectar en instancias judiciales.

3.1. Altas expectativas epistémicas

Fuera del ámbito jurídico el asunto que nos interesa ha sido discutido a propósito de la diferencia entre verdad y justificación, usualmente con el objetivo de afirmar que podemos cometer errores incluso cuando contamos con información suficiente para respaldar nuestras creencias. La falibilidad es un punto central en los debates relativos al conocimiento puesto que nos interesa saber cómo, o de acuerdo con qué criterios, podemos distinguir una decisión epistémicamente correcta de una errada[13]. En el ámbito jurídico esta discusión es igual de importante. No solo porque mantenemos la expectativa de que nuestras instituciones procesales funcionen de manera correcta, sino porque el concepto de error es uno fundante de una serie de discusiones que pueden tener lugar dentro de un proceso[14]. En general, los participantes de nuestra disciplina sostienen que se ha cometido un error con la expectativa de que "algo pase". Esperamos que los errores sean discutidos, verificados, sancionados (si corresponde), impugnados (si es posible), corregidos (si estamos a tiempo) y/o reparados (si han generado un daño). Es en este sentido en que el concepto de error es operativo: sirve a los participantes para lograr uno o varios objetivos dentro de una práctica reglada.

Conscientes de la importancia del error, los teóricos racionalistas han insistido en que este concepto opera como un presupuesto necesario para garantizar el control racional de las decisiones probatorias (Ferrer, 2005, pp. 21-22). Esta es una tesis que comparto. Solo si estamos en condiciones de sostener que los jueces pueden cometer errores, podemos también sostener que tiene sentido evaluar la corrección o incorrec-

[13] De acuerdo con Peter Winch (2011), un error un error puede definirse como una contravención de lo que se ha establecido como correcto al interior de una práctica social en relación con un conjunto de reglas o parámetros de corrección (p. 39). Si adoptamos un concepto material del error, requerimos de criterios que nos permitan identificar la verdad o falsedad de un enunciado con el que se pretende representar la realidad a la que se hace referencia. En este sentido, la dirección de ajuste de nuestras creencias determina los criterios de evaluación de un enunciado (Anscombe, 2000).

[14] Desde el punto de vista procesal, los teóricos de la prueba sostienen que la verdad de un enunciado fáctico constituye una condición necesaria, aunque no suficiente, para la correcta aplicación de una norma jurídica (Ferrajoli, 1995; Taruffo, 2002). Las razones para ello son bastante sencillas. Solo si podemos adscribir conductas realizadas a quienes efectivamente las han realizado, entonces el Derecho podrá operar como una técnica adecuada para la motivación de la conducta humana (Ferrer, 2007: 29). De esta manera, el proceso judicial se orienta, necesariamente, a la búsqueda de la verdad como objetivo prioritario (Taruffo, 2013). Una discusión sobre la tesis que sostiene la existencia de una conexión entre verdad y motivación puede consultarse en Reyes (2017).

ción de sus decisiones (Gascón, 2003, pp. 44-52). Lo que no parece ser tan intuitivo es la manera en que esta tesis (la tesis según la cual *el error es un presupuesto del control*) se relaciona con una concepción material del error probatorio. La falsedad de un juicio descriptivo difícilmente puede ser aceptada como un presupuesto para el control porque, en resumidas cuentas, no tenemos seguridad respecto de lo que realmente ha ocurrido fuera del proceso.

En el marco de esta discusión, parece especialmente relevante lo sostenido por Richard Rorty (2012, p. 12) a propósito de las diferencias que existen entre justificación y verdad. A su juicio, la principal diferencia que se da entre ambos conceptos es aquella que existe entre lo que puede ser reconocido (lo probado) y lo que no puede serlo (lo verdadero). En este sentido, sostiene que «(n)unca sabremos con seguridad que una creencia dada es verdadera, pero podemos estar seguros de que nadie es actualmente capaz de invocar objeciones residuales, de que todos coinciden en que merece ser sostenida». Siguiendo esta idea, podemos afirmar que, si no estamos en condiciones de determinar la verdad de lo que ha ocurrido, difícilmente podemos decir que los enunciados probados en el proceso coinciden o no coinciden con lo que ha ocurrido fuera de él. Y esto es un problema porque, como hemos dicho, el concepto de error debe ser útil para articular discusiones que no son solo especulativas.

Otra objeción recurrente a la utilidad de la distinción entre justificación y verdad ha sido formulada por autores que comparten una epistemología constructivista. Esta objeción apunta a que no podemos formular un juicio de correspondencia (o falta de ella) entre hechos y enunciados, entre la realidad y el lenguaje, puesto que no contamos con una unidad común que haga inteligible dicha comparación. Piense en lo siguiente. ¿Qué es lo que debemos comparar cuando comparamos enunciados y hechos? Para Torretti (2014, p. 37):

> Es evidente que el acierto o desacierto de un enunciado aseverativo no se establece comparándolo con las cosas a las que supuestamente corresponde, sino con otros enunciados que hablan de las mismas cosas con mayor proximidad y detalle, o con más autoridad. La idea misma de un cotejo directo, sin intermediación verbal, entre lo que se dice y lo que es, es oscura e implausible; pues ¿cómo podría haber una correspondencia, una adecuación entre la serie de palabras, físicamente homogéneas, ordenadas linealmente en uno o más enunciados, y la diversidad y complejidad insondables de cualquier fragmento de la realidad?

Se podría contestar a la objeción planteada que los defensores de la verdad como correspondencia en el ámbito jurídico no plantean la formulación de un juicio con estas características sino uno más uniforme: entre el enunciado que representa el hecho a probar y los enunciados que, a modo de inferencias, conectan la prueba con las generalizaciones admisibles dentro del proceso. Pero aquí el problema subsiste. Si aceptamos que los enunciados sobre hechos a probar deben compararse con otros enunciados, entonces la distinción entre lo probado y lo verdadero colapsa. Aquello que estaríamos contrastando ya no serían hechos verdaderos con hechos probados. El objeto de un juicio de contraste se produciría entre enunciados que han sido validados dentro de un dominio de conocimiento (una ciencia, una disciplina, un arte) con otros enunciados que gozan de mayores o menores credenciales epistémicas (Colo-

ma, 2017, p. 48). Si se desea seguir empleando el rótulo de verdad para designar al resultado obtenido de un juicio con estas características, probablemente nos veríamos forzados a aceptar que ya no estamos hablando de la verdad como correspondencia [15].

La apuesta de Rorty (2012, pp. 19 y ss.) es que parece sensato modificar parcialmente nuestras expectativas epistémicas o nuestras ambiciones respecto de lo que podemos lograr cuando hablamos de conocimiento. En lugar de representar estas ambiciones mediante una relación entre el sujeto que conoce y la realidad (esperando encontrar una coincidencia entre lo probado y lo verdadero), sería más sensato situar estas expectativas entre sujetos pertenecientes a diferentes comunidades discursivas. Aquí el anhelo compartido por encontrar *la verdad* podría sustituirse por el anhelo, compartido de igual forma, por obtener una *justificación adecuada* por parte de las instituciones judiciales al momento de decidir: lo que exigimos desde un punto de vista epistémico no es que un enunciado resulte ser verdadero o falso en un momento determinado, sino simplemente que esté justificado conforme a las reglas que determina un sistema de valoración probatoria [16].

Si lo dicho es correcto, entonces la distinción entre justificación (prueba) y verdad ya no sería útil para expresar la posibilidad de cometer errores cuando nuestras afirmaciones no se corresponden con el mundo. Pero todavía podría ser relevante de una manera distinta. Podríamos pensar que esta distinción es útil como un recordatorio de que, en cualquier otro momento, es posible que surjan nuevas objeciones que pongan en tela de juicio, e incluso derroten, las mejores explicaciones que han sido presentadas, pero *solo hasta ese momento*. En el ámbito jurídico, este espacio temporal se encuentra acotado en tanto nuestras discusiones procesales aspiran a consolidar un estatus normativo de manera definitiva, lo que implica que algunas explicaciones solo pueden ser desafiadas de manera excepcional, como ocurre en casos penales con el recurso de revisión. Pero este recurso tampoco exige que demostremos que el tribunal ha dicho algo falso, sino simplemente que, a la luz de los nuevos antecedentes, la explicación que ha derivado en la condena del imputado resulta insostenible. Nada excluye que, fallado el recurso, nuevos antecedentes muestren que la explicación recién anulada resulte, nuevamente, respaldada por antecedentes de mejor factura [17].

---

[15] De ser este el caso, la idea de que pueden existir errores materiales (referidos al contenido de la premisa fáctica) que no son, al mismo tiempo, errores metodológicos (referidos a la forma en que se construye la premisa fáctica) podría ser cuestionada.

[16] Nótese que esta postura no exige reemplazar una noción de la verdad como correspondencia por otra diferente pero que cumpla con una función análoga (por ejemplo, por una noción coherentista). De lo que se trata, en cambio, es de discutir la necesidad de afirmar que un enunciado es verdadero para los efectos de satisfacer nuestras expectativas epistémicas dentro del proceso.

[17] Es posible que el desajuste entre las categorías disponibles sobre los *hechos* y las expectativas epistémicas tenga su origen en la particular atención que los teóricos de la prueba han puesto en el proceso penal al momento de modelar los conceptos básicos de su disciplina (Coloma, 2024, pp. 9-11). El proceso penal ocupa un lugar central en los debates sobre la prueba porque se trata de un proceso en que las reglas sobre la suficiencia probatoria coinciden con una alta expectativa epistémica sobre los discursos que validamos en instancias judiciales. De esta forma, las exigencias probatorias permiten

Lo dicho hasta este punto nos permite abrir un espacio para discutir sobre el conocimiento de los hechos futuros como objeto de prueba, pero no como una búsqueda de verdad sino más bien como una búsqueda de justificación. La pregunta no es si podemos o no conocer el futuro (en tanto intuiciones reveladas en tiempo presentes que indiquen la verdad de lo que está por venir en el mundo). Si no, más bien, si podemos construir argumentos adecuados para sostener decisiones judiciales que presuponen consideraciones prospectivas; y si tales consideraciones pueden ser controladas racionalmente. Aquí todavía podrían subsistir objeciones relevantes de tipo epistémico en tanto no se indique con claridad cómo o de qué forma podríamos formular tales argumentos con la precisión requerida por el ordenamiento jurídico en cada caso. Para enfrentar este desafío, resulta conveniente echar un vistazo al tipo de fenómenos que nos interesa anticipar en el marco de un proceso judicial.

4. EL PROBLEMA DE HABLAR SOBRE EL FUTURO

La manera en que hablamos sobre el futuro deja en evidencia dos diferencias importantes respecto del pasado. Estas diferencias suelen conectarse con la incertidumbre, aunque probablemente el término es usado en más de un sentido. En primer lugar, decimos que "el futuro es incierto" con el objetivo de limitar nuestras expectativas sobre el conocimiento que podemos alcanzar respecto de eventos futuros. La certidumbre respecto del futuro no es común si se acepta una de las siguientes premisas: (1) que el mundo no posee una estructura determinista, de tal manera que podemos anticipar diferentes futuros o mundos posibles[18]; o bien (2) que el mundo posee una estructura determinista, pero esta es compleja, de tal manera que anticipar un evento futuro es una tarea que implica atender a la relación que existe entre múltiples factores en constante cambio (Morin, 2011, pp. 32-35). En segundo lugar, la frase suele emplearse para enfatizar que existe un margen para la acción y la planificación. A diferencia del pasado, los asuntos sobre el futuro se encuentran "abiertos", de tal manera que las acciones que realizamos en el presente condicionan los resultados que podemos llegar a esperar de manera justificada.

La diferencia en la manera en que procesamos los problemas del pasado y del futuro se encuentra también presente en la configuración histórica de nuestro lenguaje. Como señala Jouvenel (1967), en latín se reserva el vocablo *facta* para designar a los eventos pasados, mientras que se utiliza el término *futura* para hacer referencia a los eventos que todavía no han ocurrido. Solo los asuntos de *facta* pueden ser discutidos o procesados bajo los calificativos de verdadero o falso, mientras que los

---

encontrar una coincidencia entre lo probado y lo que normalmente llamaríamos "verdadero" respecto (y solo respecto) a la condena de un imputado (Dei Vecchi, 2023b, pp. 8-9). Pero estas expectativas parecen desmedidas si pretendemos satisfacerlas en procesos donde la búsqueda de la verdad ocupa un lugar secundario o cuando los hechos a probar simplemente no son susceptibles de calificarse como verdaderos o falsos en sentido estricto.

[18] En este sentido Goodman (1973), pp. 96 y ss.

asuntos de *futura* no pueden calificarse de esta manera en sentido estricto, siguiendo una noción de la verdad como correspondencia. Esto explica que la expresión "hecho futuro" (o su traducción en inglés "future fact") resulta sospechosa en tanto parece contener una contradicción en los términos siempre que hablemos de futuros contingentes; y que adoptemos una noción estricta de conocimiento en relación con la verdad como correspondencia [19].

Las diferencias idiomáticas no parecen ser especialmente relevantes para evitar o disolver esta asimetría. Como podrá anticipar el lector, la mayoría de (sino todos) los idiomas se estructuran sobre un conjunto bien definido de tiempos verbales (Cariani, 2021, p. 3), de tal manera que algunas de nuestras expresiones solo tienen sentido cuando son formuladas en relación con un estado de cosas pasado, presente o futuro. De esta forma, mientras tiene sentido afirmar que "haremos todo lo posible para evitar una tercera guerra mundial", no parece tener sentido decir que "lograremos evitar que la segunda haya tenido lugar". Las diferencias entre tiempos verbales parecen ajustarse, de esta forma, a una serie de intuiciones más profundas vinculadas con la naturaleza misma del tiempo y con la manera en que podemos justificar algunas de nuestras creencias (Torre, 2024, p. 41).

Pero hay que tener cuidado con esta afirmación. Del hecho de que nuestro lenguaje favorezca una distinción entre pasado y futuro no se sigue que el futuro sea inaccesible en términos epistémicos o que la incertidumbre respecto al pasado sea siempre menor. Es más. Así como podemos decir que nuestro lenguaje favorece la distinción entre tiempos verbales, también podemos decir que, del mismo modo, nos permite asumir cargas epistémicas sea que discutamos sobre el pasado, sobre el presente o sobre el futuro. Piense, por ejemplo, en la distinción entre "adivinación" y "pronóstico". Nadie dudaría que adivinar no es lo mismo que pronosticar porque la segunda es una actividad que se encuentra sujeta a reglas vinculadas con la manera en que podemos utilizar la información disponible. De esta manera, mientras podemos controlar un pronóstico (distinguir entre pronósticos correctos o incorrectos desde un punto de vista metodológico) no podemos hacer lo mismo con la adivinación. Quien adivina, en cambio, no se sujeta a ninguna regla (racional) sobre la manera en que se construye conocimiento o se fundamentan nuestras creencias.

En este sentido, no niego que la distinción entre hechos pasados y futuros juegue un rol relevante al momento de distribuir cargas entre las partes para los efectos de justificar una pretensión en instancias judiciales. Nuestro lenguaje favorece la distinción y no hay razones para evitarla, aunque sí para formular algunas consideracio-

[19] Ya hemos dicho, con todo, que podemos evitar esta contradicción si consideramos que el conocimiento sobre los hechos no requiere necesariamente un compromiso con la noción de la verdad como correspondencia. El compromiso epistémico adquirido al formular un enunciado fáctico en instancias judiciales no implica encontrar o buscar la verdad de lo ocurrido, sino exhibir las credenciales que legitiman nuestras afirmaciones y que permiten contrastar la afirmación con otras semejantes en prestigio, profundidad, detalle u otra cualidad similar.

nes adicionales sobre lo que significa decir que el futuro es incierto y abierto. Estas consideraciones son necesarias para precisar qué es aquello que podemos controlar en instancias judiciales cuando los jueces toman decisiones que se fundamentan en consideraciones prospectivas.

4.1. Futuro incierto y futuro abierto

Decir que el futuro es incierto parece correcto, pero también lo sería afirmar que tampoco gozamos de certidumbre cuando discutimos sobre asuntos que han acontecido en tiempo pasado (Underwood, 1979, p. 1413). Sostener una distinción tajante entre asuntos pasados y futuros puede propender a una actitud demasiado cauta respecto de las decisiones que debemos tomar en instancias judiciales. Si bien los asuntos futuros están marcados por una clara incertidumbre respecto de lo que puede llegar a ocurrir, desde un punto de vista epistémico es posible matizar esta diferencia si consideramos el conocimiento disponible al momento de tomar una decisión. En este sentido, es posible que existan situaciones en que nuestro conocimiento sobre un fenómeno futuro sea más completo que nuestro conocimiento respecto de un fenómeno presente o pasado. Si nuestro interés es justificar nuestras decisiones, entonces debiésemos considerar que el desafío de la incertidumbre es contingente a la información disponible y a nuestra comprensión del mundo, no tanto respecto al tiempo en que ha ocurrido (u ocurrirá) un fenómeno[20].

Revisemos esta idea más en detalle. En nuestro lenguaje cotidiano disponemos de diversos términos que nos permiten hacer referencia a hechos futuros, tales como predicción, pronóstico, planificación y adivinación (Decouflé, 1974)[21]. Cada uno de ellos refleja un compromiso epistémico diferente en relación con la información disponible. Si tenemos poca o nula información sobre el hecho, podemos decir que estamos adivinando. Si contamos con información completa y podemos anticipar con certeza el rumbo de los acontecimientos, podemos decir que estamos prediciendo. Si somos capaces de influir en el rumbo de los acontecimientos buscando un resultado en particular, bien podríamos sostener que estamos planificando. En cambio, si suponemos un curso de los acontecimientos basados en variables que se comportan con cierta aleatoriedad, estamos pronosticando.

[20] Curiosamente, algunos tribunales consideran que no es necesario o posible justificar sus afirmaciones sobre el futuro, quizás alentados por cierto tipo de escepticismo respecto de lo que va a ocurrir: dado que no conocemos el futuro en sentido estricto, entonces no parece necesario justificar epistémicamente nuestras consideraciones prospectivas. Algunos ejemplos de ello pueden consultarse en: Ferguson (2020) y en Dyk (2023).

[21] Estos términos denotan operaciones que son habitualmente estudiadas por disciplinas como la prospectiva. Los estudiosos del futuro a menudo sostienen que la incapacidad de obtener una certeza absoluta al formular predicciones (1) no debe entenderse como una nota distintiva de los hechos futuros (puesto que a menudo tampoco obtenemos tal certeza respecto del pasado); y (2) no justifica, por sí sola, la racionalidad de esperar a que ciertos hechos lleguen a ocurrir para los efectos de tomar una decisión.

Todos los conceptos señalados se refieren a hechos futuros, pero reflejan un compromiso epistémico diferente. Mientras que la adivinación refleja un nulo compromiso epistémico (y, por ello, se trata de una actividad incontrolable), el éxito (o el fracaso) de las restantes depende, al menos en parte, del tipo de fenómeno que estemos anticipando. De esta forma, podemos pronosticar si vendrán lluvias durante la próxima semana, predecir eclipses y planificar unas vacaciones. Parece claro que el mayor compromiso lo adquirimos al formular predicciones, pero, como veremos a continuación, los fenómenos que nos interesa predecir en el mundo jurídico no son de aquellos que realmente podemos predecir en el sentido estricto de la expresión.

Los fenómenos que pueden ser predichos cumplen con una serie de condiciones que no son habituales en el mundo físico (Staley, 2002, p. 75). Para que podamos formular predicciones es necesario que estemos ante fenómenos estacionarios, aislados y respecto de los cuales contamos con un conocimiento acabado. En su particular crítica al historicismo, Popper (1991, p. 407) argumenta que las razones que nos permiten formular predicciones en casos como el de los eclipses no deben tomarse como una base regular para formular analogías en otras áreas del conocimiento o respecto de otro tipo de fenómenos:

> Las profecías de eclipses, así como las profecías basadas en la regularidad de las estaciones (quizás las más antiguas leyes naturales comprendidas conscientemente por el hombre), sólo son posibles porque nuestro sistema solar es un sistema estacionario y repetitivo. Y esto es así debido al accidente de que se encuentra aislado de la influencia de otros sistemas mecánicos por inmensas regiones de espacio vacío, lo cual hace que esté relativamente libre de interferencias exteriores (…) Estos sistemas repetitivos son casos especiales en los que la predicción científica se hace particularmente impresionante, pero eso es todo.

Los hechos futuros que deseamos anticipar dentro del proceso judicial no cumplen con estas características. Se trata de fenómenos que no comprendemos a cabalidad y que, con frecuencia, se ven afectados por lo que ha sido denominado *paradoja del conocimiento*. Esta paradoja se produce cuando la probabilidad de que ocurran algunos eventos se encuentra directamente relacionada con nuestra ignorancia, de tal manera que, por el solo hecho de adquirir conocimiento, probablemente no ocurran de la misma forma en que lo hemos anticipado (Harari, 2024, pp. 69 y ss.). Un ejemplo claro es el funcionamiento de los mercados financieros. La sola creencia, compartida por un grupo de accionistas, de que el valor de las acciones de una empresa va a subir probablemente modifique su conducta de tal manera que las acciones terminen subiendo más de lo pronosticado inicialmente.

Es probable que los términos más adecuados para representar nuestras expectativas dentro del proceso sean los de pronóstico y planificación. Los pronósticos son especialmente útiles cuando enfrentamos fenómenos complejos cuya ocurrencia depende de múltiples factores que no podemos alterar significativamente. Este es el caso de eventos naturales como los terremotos, los tsunamis, las erupciones volcánicas y los incendios forestales. Para lidiar con la incertidumbre respecto de este tipo de eventos solemos desarrollar modelos aproximativos que pueden entregar información valiosa sobre lo que podemos llegar a esperar dentro de un rango de tiempo

(más o menos extenso). A menudo, estos modelos se construyen tomando en consideración la información disponible respecto de eventos pasados que se repiten con cierta regularidad. Luego, suponemos que un devenir normal de los acontecimientos implicará que tales eventos se repitan. Por ejemplo, sabemos que Australia enfrentará una temporada de incendios forestales cada verano incluso si no podemos anticipar con precisión el momento en que se producirá cada uno de ellos.

En ocasiones los operadores jurídicos deben tomar decisiones considerando la posible ocurrencia de eventos de este tipo para los efectos de distribuir cargas normativas vinculadas con la prevención. Así ocurrió en el caso del *Hotel Punta Piqueros*. Los tribunales nacionales debían decidir si el proyecto había cumplido con requisitos impuestos por el sistema de evaluación ambiental para los efectos de determinar su estado de avance y la posibilidad de proseguir con la construcción. Uno de los argumentos relevantes se formuló en torno a la posibilidad de que ocurriese un tsunami en la zona, lo que implicaba redoblar esfuerzos tendientes a resguardar la vida de los posibles ocupantes. En casos de este tipo, la verdad de la conjetura propuesta por el tribunal no parece ser particularmente relevante para evaluar la racionalidad de su decisión. Incluso si el tsunami no se produce (o no se produce dentro de un margen de tiempo razonable) podemos sostener que el tribunal hace bien en considerar la posibilidad de que dicho evento ocurra.

La planificación cobra especial relevancia cuando podemos incidir en la ocurrencia de un fenómeno. Frente a este tipo de eventos no nos interesa tanto saber lo que va a ocurrir sino construir un plan de acción adecuado para actuar racionalmente en el presente. Los hechos futuros de este tipo no se conciben tanto como un misterio a develar, sino que se entienden como un insumo para evaluar nuestras decisiones actuales atendiendo al futuro que queremos construir (Montfort, 2020). Las decisiones en materia cautelar pueden representarse de esta manera. Al intentar evitar un resultado disponemos de medidas preventivas que no tienen como finalidad predecir un evento futuro, sino que precisamente buscan que dicho evento, luego de ser conjeturado, no ocurra en la realidad.

Pensemos, por ejemplo, en un caso de libre competencia. Suponga que debemos determinar si la adquisición de una compañía, por parte de un competidor directo, producirá una barrera de entrada tal que impedirá a otros competidores desafiar su posición dominante dentro del mercado. Ante un caso de este tipo, se le presentan dos opciones. La primera es detener la operación y evitar la compra. Pero si toma esta medida es posible que nuestra conjetura inicial no pueda ser verificada en sentido estricto. Dado que impedimos la adquisición no estaremos en condiciones de comprobar la posible producción del monopolio que justificó nuestra medida cautelar en primer lugar. La segunda opción consiste en permitir la operación económica esperando que el monopolio simplemente no llegue a producirse. Y aquí podemos obtener dos resultados. Que el monopolio no se forme, caso en que podríamos pensar que la medida cautelar no era necesaria. O que el monopolio se forme, caso en que nuestra conjetura podría considerarse comprobada: el resultado que habíamos

conjeturado finalmente se ha producido en la realidad. Nuestra conjetura ha resultado ser verdadera, pero, si este es el caso, parece contraintuitivo afirmar que el juez actuó de forma correcta. Ya sea que estemos pronosticando o planificando, la verdad de nuestra conjetura, por sí sola, no parece ser un dato relevante para evaluar nuestras decisiones dentro del margen temporal del proceso judicial.

Entender los enunciados sobre hechos futuros como planificaciones y pronósticos (y no como predicciones) es relevante porque nos permite identificar los criterios necesarios para evaluar la calidad de una decisión. Al formular una predicción nos comprometemos con la posibilidad de comprobar o verificar la verdad de nuestro enunciado, un compromiso similar al que, según los teóricos de la prueba, adquirimos al probar hechos pasados. En este sentido, las predicciones fracasan (son erradas) cuando existe una discrepancia entre lo aventurado y lo que realmente ha ocurrido. Se trata de un juicio *a posteriori* que tiene sentido luego de que podemos referenciar un resultado que opera como criterio de corrección. Cuando dicho resultado coincide con el valor anticipado, podemos sostener que nuestra predicción ha sido exitosa[22].

Los pronósticos y las planificaciones no pueden evaluarse de esta manera porque nuestro compromiso inicial no guarda una relación estrecha con la verdad o falsedad del enunciado. En el caso de los pronósticos ello ocurre porque se emplean modalizadores del discurso (como "probablemente") que, *grosso modo*, son compatibles con la posibilidad de que el evento conjeturado no ocurra de la manera en que hemos anticipado. La planificación, en cambio, nos compromete con un modo de actuar que resulta adecuado para obtener un resultado que se estima valioso, pero el juicio de corrección es *a priori*. Esto es así porque muchas veces planificamos con la expectativa de evitar que un resultado se verifique en la realidad. Que el hecho conjeturado no llegue a ocurrir no implica, en estos casos, que la decisión sea incorrecta (como sí parece ocurrir en el caso de la predicción). Todavía más, podemos imaginar casos en que el éxito de la planificación coincida con la no ocurrencia del fenómeno que ha sido proyectado[23].

4.2. Futuro y error

Si entendemos que los enunciados sobre hechos futuros denotan operaciones de pronóstico o planificación, no tiene sentido procesal preguntarnos si el hecho conjeturado ha resultado ser verdadero o falso. En el caso de los pronósticos, la pregunta resulta poco relevante porque la conjetura podría ser correcta incluso si el hecho que

[22] Aquí estoy suponiendo que la verificación o comprobación es posible en el caso de los enunciados sobre hechos pasados, pero como ya hemos señalado, existen razones que nos permiten dudar de que tal ejercicio de comprobación tenga sentido en instancias judiciales.

[23] Esto, por ejemplo, podría ocurrir en casos de justicia cautelar, donde el éxito del proceso depende de evitar que un resultado se produzca en la realidad.

hemos anticipado finalmente no se verifica en la realidad. Algo similar ocurre en el caso de las planificaciones, en donde nuestra decisión puede tener por finalidad lograr que el evento finalmente no ocurra. Aquí la no ocurrencia del hecho conjeturado podría incluso operar como un indicativo del éxito judicial y no de la existencia de un error.

Si entendemos que los enunciados sobre hechos futuros denotan operaciones de predicción en sentido estricto, tampoco tiene sentido procesal la pregunta por la falsedad del hecho. Esto porque dicha verificación solo es posible *a posteriori*, luego de que hemos tomado una decisión y podemos contrastar el contenido de nuestro enunciado inicial con un resultado comprobable en la realidad. Ello implica que nuestra verificación ocurre luego de que el proceso ha alcanzado el estatus de cosa juzgada material. Y ello es importante porque entonces, si se ha cometido un error, se trata de un error que no tiene la aptitud para modificar el resultado del proceso. Será un error que no posee relevancia jurídica.

Lo dicho no implica que debamos renunciar al control racional de las decisiones judiciales sobre hechos futuros. Tanto los pronósticos como las planificaciones pueden ser controladas *a priori*, antes de que ocurran (si es que ocurren) los eventos a los que hacen referencia. Pero para garantizar dicho control, parece necesario ajustar nuestras expectativas epistémicas identificando qué podemos decir sobre el futuro con la seguridad suficiente como para justificar una decisión. Para ello, parece especialmente relevante atender al tipo de proceso judicial en que se debe realizar la consideración prospectiva. Dentro de un proceso judicial civil por daño futuro o en un proceso constitucional ante la amenaza a un derecho fundamental nuestras conjeturas parecen representarse, de mejor manera, mediante conceptos como los de planificación y pronóstico. Ello nos permite identificar criterios o marcadores del error judicial que no requieren afirmar que un enunciado es verdadero o falso en sentido estricto, siguiendo una noción de la verdad como correspondencia.

5. CONCLUSIONES Y ALGUNOS DESAFÍOS PENDIENTES

El estudio de los hechos futuros presenta una serie de desafíos desde un punto de vista lógico, semántico y epistémico. Desde el punto de vista judicial, con todo, las tareas pendientes parecen ser dos. La primera de ellas consiste en ajustar nuestras expectativas epistémicas al tipo de hecho futuro que debemos conjeturar dentro de los procesos judiciales. Ya sabemos que predecir hechos futuros es una tarea compleja que solo parece posible respecto de eventos sobre los cuales tenemos un conocimiento acabado y que acontecen de manera regular en el mundo físico, como el movimiento de los cuerpos celestes o la subida de la marea. Respecto de los eventos que habitualmente debemos anticipar en los procesos judiciales este no es el caso, de tal manera que las operaciones que realizamos no son predicciones en sentido estricto: no podemos predecir el comportamiento de los mercados financieros, el desarrollo futuro de un niño o la fecha exacta en que ocurrirá un tsunami. Esto es importante

porque distribuimos cargas y beneficios entre ciudadanos que se fundan en razonamientos prospectivos. Entregamos a uno de los padres la tuición pensando en que el menor se desarrollará de mejor manera; exigimos mayores medidas preventivas a un hotel que se instalará en el borde costero; y evitamos la compra de un actor económico por parte de un competidor directo porque pensamos que ello generará un monopolio. Pensar los hechos futuros desde la posibilidad de formular predicciones contribuye a que los operadores jurídicos fijen exigencias desmedidas a las partes al momento de justificar sus pretensiones.

Para ajustar nuestras expectativas epistémicas conviene desarrollar un vocabulario que nos permita distinguir entre diferentes operaciones mediante las que procesamos los problemas sobre el futuro. Los prospectivistas distinguen entre adivinanzas, pronósticos, planificaciones y predicciones. Cada una de estas operaciones media entre las posibilidades de actuar y las posibilidades de, en un momento futuro, verificar o comprobar. Pero el foco de los estudios sobre el futuro se ubica en las posibilidades de actuar para construir un futuro deseable, no tanto en la posibilidad de predecir un futuro con la pretensión de formular enunciados que, a la postre, resulten verdaderos. Los procesos judiciales en que decidimos hechos futuros suelen estructurarse de la misma forma: el foco de interés está en evitar resultados inaceptables mediante acciones en tiempo presente. De esta forma, el futuro que interesa en un sentido procesal es el futuro que puede construirse mediante las decisiones adoptadas por el juez.

La segunda tarea pendiente consiste en definir las condiciones del error judicial en el caso de los hechos futuros. Uno de los avances más notorios de la teoría del razonamiento probatorio consiste en el intento por disciplinar las decisiones fácticas de los tribunales de justicia a partir de la distinción entre hechos probados y hechos verdaderos. El problema aquí es doble. Por un lado, los operadores jurídicos no tienen acceso a los segundos como una entidad independiente de los primeros. Lo que podemos considerar para tomar decisiones correctas son los hechos probados, no los hechos que realmente han acontecido fuera del proceso. Por otro, los enunciados que se validan en procesos judiciales no siempre son susceptibles de ser calificados como verdaderos o falsos siguiendo una noción de la verdad como correspondencia.

La importancia de definir los criterios de corrección que podemos emplear para evaluar decisiones sobre hechos futuros radica en la posibilidad de controlar la actividad judicial y uniformar la administración de justicia. Esto ocurre porque las diversas operaciones que posibilitan el control de las decisiones judiciales requieren que estemos en condiciones de sostener que el juez ha cometido un error significativo. Este es el caso de operaciones como la apelación o la casación. Pero ocurre que la verdad como correspondencia no es un buen criterio. En el caso de los hechos futuros el criterio de la verdad como correspondencia no es adecuado porque tales enunciados no pueden calificarse como verdaderos o falsos dentro de los confines del proceso judicial. Esto exige desarrollar un esfuerzo doctrinal en construir criterios alternativos que puedan ser útiles para evaluar las conjeturas propuestas por los tribunales de justicia.

BIBLIOGRAFÍA

Accatino, D. (2019). ¿Somos todos racionalistas ahora? *Revus: Journal for constitutional theory and philosophy of law*, 39(1), pp. 85-102.

Amsterdam, A. and Bruner, J. (2002). *Minding the Law*. Harvard University Press.

Anscombe, E. (2000). *Intention*. Harvard University Press.

Austin, J. (1971). *Cómo hacer cosas con palabras*. Paidós.

Bayón, J. (2008). Epistemología, moral y prueba de los hechos: hacia un enfoque no Benthamiano. *Analisi e Diritto*, 2(4), pp. 15-34.

Bravo, S. (2023). Lenguaje, lógica y algunas repercusiones de la (in)definición del thema probandum. *Quaestio facti: Revista Internacional de Derecho Probatorio*, 4(1), pp. 61-94.

Briggs, R. and Forbes, G. (2011). The Real Truth About the Unreal Future. In K. Bennett (ed.), *Oxford Studies in Metaphysics volume 7* (pp. 257-304), Oxford University Press.

Carbonell, F. (2022). Un modelo de decisión judicial justificada para el proceso penal chileno. *Política Criminal*, 17(33), pp. 58-84.

Carbonell, F. (2023). Los errores del juez. Presupuestos y tipologías, en J. García Amado (ed.), *El error judicial. Problemas y regulaciones* (pp. 57-86). Tirant Lo Blanch.

Cariani, F. (2021) *The Modal Future*. Cambridge University Press.

Coloma, R. (2016). Las disciplinas jurídicas y su reinvención. *Revista Ius et Praxis*, 22(2), pp. 253-298.

Coloma, R. (2017). Bases de un modelo conceptual para decidir hechos. *Doxa: Cuadernos de filosofía del derecho*, 40(1), pp. 69-92.

Decouflé, A. (1974). *La prospectiva*. Oikos-tau.

Dei Vecchi, D. (2020). Prueba libre, justificación epistémica y el noble sueño de los estándares de prueba. *Revista de Derecho (Valdivia)*, 33(2), pp. 25-48.

Dei Vecchi, D. (2023a). Sentencia judicial, prueba y error: El rol de la verdad de las premisas fácticas en la aplicación de normas jurídicas y en la justificación de decisiones judiciales. *Isonomía: Revista de teoría y filosofía del derecho*, 58(1), pp. 107-148.

Dei Vecchi, D. (2023b). Sentencias judiciales con premisas falsas: entre fallas, errores y exageraciones. *Revus. Journal for Constitutional Theory and Philosophy of Law*, 49(1).

Dyk, T. (2023). The Role of Non-Adjudicative Facts in Judicial Decision Making. *Stanford Law Review Online*, 76(1), pp. 10-29.

Ferguson, B. (2020). Predictive Facts. *Washington Law Review*, 95(3), pp. 1621-1672.

Ferrajoli, L. (1995). *Derecho y razón. Teoría del garantismo penal*. Trotta.

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

Ferrer, J. (2007). *La valoración racional de la prueba* (2ª edición). Marcial Pons.

Ferrer, J. (2018). Prolegómenos para una teoría sobre los estándares de prueba, en D. Papayannis y E. Pereira (eds), *Filosofía del derecho privado* (pp. 401-430). Marcial Pons.

Ferrer, J. (2021). Proven facts, Beliefs, and Reasoned Verdicts, in C. Dahlman, A. Stein, y G. Tuzet (eds), *Evidence Law* (pp. 40-52). Oxford University Press.

Gascón, M. (2003). Concepciones de la prueba: Observación a propósito de Algunas consideraciones sobre la relación entre prueba y verdad, de Michele Taruffo. *Revista Discusiones*, 3(1), pp. 43-54.

Gascón, M. (2005). Sobre la posibilidad de formular estándares de prueba objetivos. *Doxa: Cuadernos de filosofía del derecho*, 28(1), pp. 127-139.

Gascón, M. (2010). *Los hechos en el derecho* (3ª edición). Marcial Pons.

Gascón, M. (1999). La naturaleza prealgebraica de la matemática escolar. *Educación Matemática,* 11(1) pp. 77-88.

Giannini, H. (2021). *Desde las palabras: A propósito de la rectitud de los nombres*. Ediciones Universidad Diego Portales.

González Lagier, D. (2022). *Quaestio facti: Nuevos ensayos sobre prueba y filosofía. Volumen I* (2ª edición). Palestra.

Harari, Y. (2024). *Homo Deus: Breve historia del mañana*. Debate.
Hunter, I. (2015). Las dificultades probatorias en el proceso civil. *Revista de Derecho de la Universidad Católica del Norte*, 22(1), pp. 209-257.
Jouvenel, B. (1969). *The Art of Conjecture*. Basics Books.
Laudan, L. (2013). *Verdad, error y proceso penal*. Marcial Pons.
MacFarlane, J. (2003). Future Contingents and Relative Truth. *The Philosophical Quarterly*, 53(212), pp. 321-336.
Marchese, E. (2019). De incertidumbre no se muere. *Eunomía. Revista de Cultura de la legalidad*, 17(1), pp. 8-34.
Montfort, N. (2020). *El futuro*. Ediciones Universidad Católica de Chile.
Mora, J. (2021). Predictibilidad conductual y proceso penal. Algunos apuntes sobre el fundamento epistémico de los hechos futuros en las medidas cautelares. *Quaestio Facti*: *Revista Internacional de Derecho Probatorio*, 2(1), pp. 53-85.
Morin, E. (2011). Introducción al pensamiento complejo. Gedisa.
Muñoz Sabaté, L. (2001). *Fundamentos de prueba judicial.* J. M. Bosch.
Øhrstrøm, P. and Hasle, P. (2020). *Future Contingents*. The Stanford Encyclopedia of philosophy. Disponible en: *https://plato.stanford.edu/archives/win2019/entries/future-contingents*.
Popper, K. (1991). *Conjeturas y refutaciones*. Paidós.
Reyes, S. (2017). Sobre derecho y la averiguación de la verdad. *Doxa: Cuadernos de Filosofía del Derecho*, 40(1), pp. 317-336.
Rivera, R. (2011). *La prueba: un análisis racional y práctico*. Marcial Pons.
Rorty, R. (2012). *Sobre la Verdad: ¿validez universal o justificación?* Amorrortu Editores.
Russell, B. (2018). *Ensayos filosóficos* (3ª edición). Alianza.
Staley, D. (2002). A History of the Future. *History and Theory*, 41(4), pp. 72-89.
Taruffo, M. (2002). *La prueba de los hechos* (4ª edición). Trotta.
Taruffo, M. (2013). La verdad en el proceso. *Derecho y sociedad*, 40(1), pp. 239-248.
Torre, S. (2011). The Open Future. *Philosophy Compass*, 6(5), pp. 360–373.
Torre, S. (2024). Knowledge of the Future and Reliable Belief Forming Processes. *Australasian Journal of Philosophy*, 130 (1), pp. 39-57.
Torretti, R. (2014). *Estudios filosóficos 2011-2014.* Ediciones Universidad Diego Portales.
Tuzet, G. (2021). *Filosofía de la prueba jurídica*. Marcial Pons.
Underwood, B. (1979). Law and the Crystal Ball: Predicting Behavior with Statistical Inference and Individualized Judgment. *Yale Law Journal*, 88(7), pp. 1408-1449.
Uviller, R. (1982). Evidence of character to prove conduct. Illusion, Illogic, and injustice in the courtroom. *Penn Law Review*, 130(1), pp. 845-891.
Valenzuela, J. (2022). Sin convicciones sobre el futuro. Una observación a "Prueba sin convicción: estándares de prueba y debido proceso" de Jordi Ferrer a propósito de la prueba cautelar. *Revista Discusiones*, 28(1), pp. 159-177.
Winch, P. (1958). *Ciencia social y filosofía*. Amorrortu editores.

Quaestio facti. Revista Internacional sobre Razonamiento Probatorio
Quaestio facti. International Journal on Evidential Legal Reasoning
Sección: Ensayos
2026 I 10 pp. 245-282
Madrid, 2025
DOI: 10.33115/udg_bib/qf.i10.23191
Marcial Pons Ediciones Jurídicas y Sociales
© Mauricio Duce J.
ISSN: 2604-6202
Recibido: 30/06/2025 | Aceptado: 04/11/2025 | Publicado *online*: 01/12/2025
Editado bajo licencia Reconocimiento 4.0 Internacional de Creative Commons

LOS INFORMES PERICIALES ESCRITOS EN EL PROCESO PENAL EN CHILE: ANÁLISIS CRÍTICO DE SU REGULACIÓN Y PRÁCTICA*

Mauricio Duce J.
Escuela de Gobierno Universidad Católica de Chile
mauricio.duce@uc.cl

RESUMEN: El trabajo analiza críticamente, a través de un estudio dogmático y empírico exploratorio, la regulación legal y práctica en Chile en materia de informes periciales escritos en el proceso penal. El texto destaca su importancia para facilitar un escrutinio efectivo de esta prueba en los sistemas acusatorios y luego revisa estándares comparados y ejemplos de derecho extranjero sobre contenidos mínimos que debiera incluir para cumplir adecuadamente tal función. A continuación, se analiza críticamente la regulación del Código Procesal Penal de Chile a la luz de estos estándares y ejemplos, destacando su insuficiencia, y luego se analizan los contenidos que en la práctica habitual del sistema tienen estos documentos. Para esto se utilizan los escasos antecedentes empíricos disponibles en estudios previos y se muestran los resultados de una investigación exploratoria sobre la base de noventa informes periciales solicitados por el Ministerio Público (n: 33) y la Defensoría Penal Pública (n: 57).

PALABRAS CLAVE: prueba pericial; informe pericial escrito; proceso penal; control probatorio; contenidos mínimos.

* Trabajo desarrollado en el proyecto de investigación Fondecyt Regular n° 1240077 ¨Investigación dogmática sobre herramientas procesales para lidiar con los sesgos cognitivos de los peritos en el ámbito penal¨ en el que el autor tiene la calidad de investigador responsable. También ha contado con el apoyo de ANID -Programa de investigación científica Milenio- NCS2024_058, en el que el autor es investigador principal. Se agradece la colaboración de Angélica Torres, de la Facultad de Derecho de la Universidad Diego Portales, en la búsqueda jurisprudencial y edición final del texto.

EXPERT REPORTS IN CRIMINAL PROCEDURE IN CHILE: A CRITICAL ANALYSIS OF THEIR REGULATION AND PRACTICE

ABSTRACT: The paper critically analyzes, through a dogmatic and exploratory empirical study, the legal and practical regulation in Chile regarding written expert reports in the criminal procedure. The text highlights the importance of such documents to facilitate an effective scrutiny of this evidence in accusatory systems and then reviews comparative standards and examples of foreign law on minimum contents that should be included to adequately fulfill such function. The regulation of the Chilean Code of Criminal Procedure is critically analyzed in the light of these standards and examples, highlighting its inadequacy, and then the contents of these documents in the usual practice of the system are analyzed. For this purpose, it is reviewed the limited empirical evidence available in previous studies and the results of exploratory research based on 90 expert reports requested by the Public Prosecutor's Office (n: 33) and the Public Defender's Office (n: 57).

KEYWORDS: expert evidence; written expert reports; criminal procedure; evidence control, minimum content.

SUMARIO: 1. INTRODUCCIÓN.— 2. LA FUNCIÓN DEL INFORME PERICIAL ESCRITO Y SU REGULACIÓN EN EJEMPLOS DE DERECHO EXTRANJERO: 2.1. La función básica del informe escrito; 2.2. Estándares, propuestas y regulaciones de derecho extranjero en materia de informes periciales escritos: 2.2.1. Contenidos básicos del informe y exigencias documentación y registro en su elaboración; 2.2.2. Información relevante del experto; 2.2.3. Oportunidad y acceso al informe y sus archivos o registros.— 3. LA REGULACIÓN LEGAL DEL INFORME PERICIAL EN EL CPP EN CHILE: 3.1. Los contenidos del informe pericial escrito (artículo 315 del CPP); 3.2. Los antecedentes del experto (artículo 314 del CPP); 3.3. Oportunidad en el acceso al informe escrito y los antecedentes del experto.— 4. LOS INFORMES PERICIALES ESCRITOS EN LA PRÁCTICA DEL SISTEMA DE JUSTICIA CRIMINAL CHILENO: 4.1. La evidencia disponible; 4.2. Resultados de una investigación empírica exploratoria: 4.2.1. Caracterización general de los informes revisados; 4.2.2. Principales hallazgos en materia de contenidos: 4.2.2.1. Heterogeneidad de los informes; 4.2.2.2. Informes incompletos.— 5. A MODO DE CONCLUSIÓN.— BIBLIOGRAFÍA.

1. INTRODUCCIÓN

La prueba pericial ha adquirido una creciente relevancia en los sistemas de justicia penal contemporáneos y Chile no es una excepción. Por ejemplo, un estudio muestra que un 83,1% de los jueces penales y un 92% de los abogados encuestados consideró que esta prueba era relevante o muy relevante al momento de dictar sentencia [Dirección de Estudios de la Corte Suprema (DECS), 2017, p. 34]. Ideas similares han sido expresadas en diversas investigaciones[1] o pueden concluirse a partir de los datos

[1] Sobre el uso de la prueba pericial en general en los procesos penales (Duce, 2018, pp. 51-59). Referido a la importancia de los peritajes pediátricos forenses en la investigación de muertes infantiles (Fundación Amparo y Justicia, 2024, p. 41). Destacando su rol clave como insumo para las defensas

disponibles sobre la producción y el uso de esta evidencia en los procesos penales chilenos[2].

A pesar de esto, se trata de una materia que no ha recibido la atención requerida de parte de la doctrina chilena, especialmente en investigaciones empíricas destinadas a evaluar su uso en la práctica. Un aspecto en que el estudio es particularmente escaso es el referido al informe pericial escrito y sus contenidos.

En el sistema procesal penal acusatorio implementado a partir del año 2.000, la prueba pericial es una evidencia compleja que se integra con dos componentes[3]. En el primero, la evidencia que puede ser valorada por el tribunal es, por regla general, la declaración prestada por el o la perito en la audiencia de juicio[4]. En el segundo, para que esa declaración sea posible, se requiere que en forma previa se haya elabora-

---

penales y su impacto aumentando las absoluciones en juicio (Departamento de Estudios Defensoría Nacional, 2016, pp. 21-23, 30). En el contexto del uso de peritajes psicológicos en el sistema (Condemarín y Lara, 2015, p. 4).

[2] Considerando a cuatro de las principales instituciones estatales que elaboran informes periciales se habrían producido más de 230 mil en el año 2023. Así, el Servicio Médico Legal (SML) reporta la elaboración de 146.853 informes (no todos para el ámbito penal), el Laboratorio de Carabineros de Chile (LABOCAR) 33.745, el Laboratorio de Criminalística de la Policía de Investigaciones de Chile (LABOCRIM) 29.053 y el Instituto de Salud Pública 24.778 análisis de composición de sustancias ilícitas (*https://www.ispch.cl/wp-content/uploads/2024/05/Documento_Cuenta_Publica.pdf*). Por su parte, la Defensoría Penal Pública reporta en el año 2024 que encargaron y financiaron 15.898 informes periciales por un monto superior a los 1.900 millones de pesos (alrededor de dos millones de dólares de los Estados Unidos) (Defensoría Penal Pública, 2025, p. 33). Por otra parte, una investigación publicada en 2017 identificó que los fiscales ofrecían prueba pericial a juicio en un 28% de sus acusaciones (Arellano, 2017, p. 65).

[3] Reconociendo la existencia de estos dos componentes, a modo ejemplar, pueden verse las sentencias de la Corte Suprema Rol n° 4554-2014 de 10 de abril de 2014 y Rol n° 16.613-2024 de 19 de julio de 2024; de la Corte de Apelaciones de Iquique Rol n° 78-2004 de 26 de noviembre de 2004; y, de la Corte de Apelaciones de Antofagasta Rol n° 877-2023 de 3 de julio de 2023.

[4] De conformidad al artículo 329 del Código Procesal Penal (en adelante CPP) la declaración personal del perito en el juicio no puede sus sustituida por la lectura de registros u otros documentos en donde ellas consten. Excepciones a esta regla se regulan en el inciso final del mismo artículo 329 y el 331. Los tribunales superiores de Chile se han pronunciado en reiteradas ocasiones y en forma bastante inequívoca afirmando esta regla general, anulando casos en que ella no ha sido cumplida. Entre otras, Corte Suprema Rol n° 5597-2009 de 2 de noviembre de 2009; Corte Suprema Rol n° 14553-2024 de 15 de julio de 2024; Corte de Apelaciones de Valparaíso Rol n° 865-2010 de 8 de octubre de 2010; Corte de Apelaciones de Valdivia Rol n° 263-2005 de 2 de enero de 2005; Corte de Apelaciones de Iquique Rol n° 78-2004 de 26 de noviembre de 2004. En un caso reciente, la Corte Suprema rechazó un recurso de nulidad interpuesto por el defensa fundado en vulneraciones al derecho a confrontación al haberse permitido el reemplazo de la comparecencia del experto por la lectura de su informe estando en una situación de imposibilidad de asistencia. La Corte rechaza el recurso al considerar que, en el caso específico, no se trató de una infracción sustancial ya que la convicción condenatoria del tribunal surgió de otra prueba presentada en el juicio. Corte Suprema Rol n° 246.257-2023 de 9 de enero de 2024. La Corte Suprema también ha precisado que la regla general del artículo 329 no se hace extensible a procesos de extradición pasiva. Corte Suprema rol n° 97.049-2021 de 29 de diciembre de 2021. Un ejemplo de sentencia que analiza el alcance de las excepciones a la comparecencia reguladas en el artículo 315 del CPP: Corte de Apelaciones de Antofagasta Rol n° 877-2023 de 3 de julio de 2023.

do un informe escrito de parte del experto que declara, el que debe incluir diversos contenidos[5]. Quizás esta estructura compleja ha llevado a que a la fecha se haya prestado más atención a lo que ocurre en la audiencia de juicio que en la elaboración del informe escrito, el que en principio no sería el medio de prueba.

Este trabajo pretende ayudar a superar esa brecha de conocimiento por medio de exponer los resultados de una investigación dogmática y empírica exploratoria sobre la normativa y práctica de los informes escritos en materia penal en Chile. Se sostendrá que la regulación del CPP es insuficiente para satisfacer las necesidades de información que se requieren en la actualidad para lidiar adecuadamente con esta prueba. Se mostrará también evidencia que indica que, en la práctica, los informes escritos presentan varios problemas respecto a sus contenidos que dificultan que dicho documento favorezca un control adecuado de esta prueba. El examen de algunos ejemplos de derecho extranjero y propuestas y estándares disponibles en la materia mostrará, en cambio, que sobre este punto existe un importante desarrollo comparado, el que se ha traducido en la regulación de reglas mucho más completas y precisas respecto a los contenidos mínimos que debiera tener el informe pericial escrito.

El trabajo se desarrollará en tres apartados además de esta introducción y las conclusiones. En el segundo, se analizarán brevemente las funciones que cumple el informe escrito en el proceso penal y se mostrarán algunos ejemplos seleccionados del derecho extranjero sobre su regulación y contenidos. En el tercero se realizará un análisis crítico de la normativa vigente en Chile contrastándola con los ejemplos revisados en la sección previa y con algunas opiniones expresadas por la doctrina y la jurisprudencia. En el cuarto, se expondrá evidencia que sugiere que la práctica actual se caracteriza por el uso de informes periciales escritos con niveles muy diversos de desarrollo y completitud, dificultando que este instrumento cumpla con la función de favorecer un escrutinio estricto de esta evidencia. Para estos efectos, el texto se basará principalmente en los resultados de una investigación empírica exploratoria realizada sobre la base del análisis de contenidos de 90 informes periciales escritos producidos a requerimiento del Ministerio Público (en adelante MP) (n: 33) y la Defensoría Penal Pública (en adelante DPP) (n: 57) de Chile. En las conclusiones se presentarán algunas breves reflexiones sobre la necesidad de avanzar en cambios de práctica y de regulación legal.

---

[5] El artículo 315 del CPP señala que este informe se debe realizar sin perjuicio de la comparecencia del perito al tribunal y debe contener diversos elementos que se analizarán más delante de manera específica. Sobre esta materia, también existen pronunciamientos de tribunales superiores rechazando la admisibilidad de prueba pericial a juicio que no cuenta con soporte en un informe escrito. Corte de Apelaciones de Chillán Rol nº 93-2007 del 8 de agosto de 2007. Existe una posición doctrinaria en el país que, reconociendo el carácter complejo de la prueba pericial, sostiene que en juicio oral debe recibirse tanto la declaración del perito como su informe escrito (Cruz, 2023, pp. 315-316; Maturana y Montero, 2010, p. 1.042). Una evaluación del sistema acusatorio de 2017 identificó que en una de las jurisdicciones observadas (Valparaíso) se detectó la práctica de presentar los informes periciales escritos como prueba documental aun cuando el perito compareció y prestó declaración en juicio. El texto critica esta situación, pero reconoce sería excepcional (Arellano, 2017, p. 177).

2. LA FUNCIÓN DEL INFORME PERICIAL ESCRITO Y SU REGULACIÓN EN EJEMPLOS DE DERECHO EXTRANJERO

En este apartado se analizarán las funciones que cumple el informe escrito. Luego, se revisarán algunos ejemplos seleccionados de derecho extranjero que lo regulan y algunas propuestas existentes en el ámbito comparado sobre ese aspecto.

2.1. La función básica del informe escrito

En un contexto procesal en el que la prueba valorable es la que se presenta en la audiencia de juicio a través del testimonio del experto, una cuestión básica es comprender las razones que justifican la exigencia de un informe escrito previo o la función que este cumpliría. Se trata de un punto en que pareciera existir un nivel de consenso importante, que puede traducirse a la idea que sin informe no es posible ni realista controlar adecuadamente la calidad y confiabilidad de esta evidencia[6]. Los informes escritos cumplirían entonces con la función de mejorar las posibilidades para su control o escrutinio[7]. Desde esta perspectiva, su existencia ha sido asociada a exigencias de debido proceso, al ejercicio del derecho a confrontación y a necesidades epistémicas de los sistemas de justicia penal para adoptar decisiones acertadas (Vázquez, 2022, pp. 5-7; Vázquez, 2024, p. 253; Meza, 2024, p. 294; Urrea, 2024, p. 83).

Como ha destacado un autor, este análisis no debe olvidar que el informe escrito es normalmente utilizado como fuente de información para adoptar diversas decisiones antes del desarrollo del juicio oral (González-Wilhelm, 2018, pp. 342-344). Estas pueden ser de persecución penal o jurisdiccionales, por ejemplo, de medidas cautelares como la prisión preventiva. En el caso de ciertos procedimientos especiales, como el abreviado, incluso la prueba a valorar por el tribunal será el informe escrito. En todas estas situaciones la prueba pericial es el informe escrito y de ahí también su enorme relevancia. Como ha expresado la National Commission on Forensic Science (en adelante NCFS) de los Estados Unidos, en estos casos la función de los informes es crítica y, por lo mismo, requiere que sean redactados con claridad y precisión, conteniendo información suficiente que sirva para la adopción de decisiones en el sistema de justicia criminal (NCFS, 2015).

---

[6] Por ejemplo, investigación reciente que muestra que los peritos están expuestos a sesgos cognitivos que podrían afectar la calidad y confiabilidad de sus resultados concluye que un elemento crucial para identificar su existencia y reducir su impacto es por medio de contar con acceso a información precisa respecto a cómo fue construido el peritaje y se arribó a sus conclusiones (Neal *et al.*, 2022, p. 179). En general sobre la importancia de los informes escritos para el control de esta prueba (Summersby *et al.*, 2024).

[7] En la doctrina chilena se trata de una idea frecuentemente mencionada como la justificación de la exigencia de informe escrito previo del artículo 315 del CPP (Aguilera y Galleguillos, 2025, pp. 22-24; Cerda, 2010, p. 457; DESC, 2017, pp. 10-11; Duce, 2013, pp. 148-154; González-Wilhelm, 2018, pp. 344-347; Horvitz y López, 2004, pp. 297-298). También ha sido recogida por la jurisprudencia de la Corte Suprema (Fierro y Walker, 2024, pp. 358-361).

La necesidad de contar con dispositivos que mejoren la capacidad de control de la prueba pericial surge debido a la naturaleza especial de esta evidencia: el aportar un conocimiento experto (científico, técnico o de otra naturaleza) que está fuera del conocimiento de los juzgadores y operadores legales de la justicia penal. El que ellos deban valorar o tomar decisiones sobre información en principio extraña a su formación y conocimiento expone al sistema a distintos riesgos. Estos incluyen la posibilidad que en la práctica (de facto) se delegue la toma de decisiones en los peritos[8]; se sobrevalore su contenido (National Research Council of the National Academies of Science (NRC-NAS), 2009, pp. 185-186; Summersby *et al.*, 2024, pp. 1-2) o se la valore en forma incorrecta produciendo decisiones erróneas[9]; o esto se haga a través de un procedimiento y con información que no permita una evaluación racional de la misma (Edmond, 2013; Edmond, 2015; Vázquez, 2022, p. 7).

En ese escenario, la doctrina chilena ha destacado que el informe escrito previo contribuye de diversas maneras a mejorar el escrutinio de esta prueba. Una primera se produce al facilitar el conocimiento a las partes sobre el peritaje realizado de manera de estar en condiciones para preparar debates de admisibilidad (eventualmente obtener la exclusión) de esta prueba en el evento que ella no satisfaga las exigencias establecidas en la legislación para ingresar a juicio (Duce, 2013, pp. 150-151; González-Wilhelm, 2018, pp. 344-345)[10]. Admitido a juicio, el acceso oportuno al informe escrito es la pieza clave que permite a los litigantes preparar su presentación en el examen directo (Blanco et al, 2021, p. 222)[11] y su confrontación por medio del contra examen (Aguilera y Galleguillos, 2025, p. 26; Duce, 2013, pp. 151-152; González-Wilhelm, 2018, pp. 345-346). En juicio, el informe cumple también un rol de control del contenido de la declaración pudiendo usarse como una herramienta que permite evidenciar contradicciones entre sus contenidos y las declaraciones en audiencia[12] o fijando los límites sobre los cuáles puede referirse el experto en su

[8] Se trata de una preocupación presente en la literatura comparada desde los inicios del siglo XIX que se basa en el riesgo de una excesiva deferencia a la opinión de los expertos (Appazov, 2016, p. 22). La escasa evidencia empírica disponible en Chile sugiere que se trataría de un riesgo real (DECS, 2017, pp. 37-38; Duce, 2018, pp. 81-85; Urrea, 2024, p. 274).

[9] Sobre el vínculo entre uso de prueba pericial y condenas erróneas existe bastante literatura, recomiendo como punto de partida un texto clásico (Garrett, 2011, pp. 84-117) y una investigación más reciente (Morgan, 2023a). En español y con evidencia de Chile (Duce, 2022, pp. 143-183).

[10] La Law Commission de Inglaterra y Gales en su informe de 2011 recomendó elevar los estándares de descubrimiento de la prueba pericial en dicho país al constatar que era la única forma de permitir un control de admisibilidad más estricto (The Law Commission, 2011, párr. 7.35, p. 116). The Law Commission es un cuerpo independiente establecido en 1965 que tiene por propósito hacer una revisión sistemática de las leyes de Inglaterra y Gales y, eventualmente, proponer reformas y mejoras: *https://lordslibrary.parliament.uk/law-commission-bills-codifying-the-law-and-tidying-up-the-statute-book/*

[11] Estos autores destacan el potencial uso del informe escrito para refrescar la memoria del perito frente a un olvido o laguna en su examen directo en juicio de acuerdo con lo previsto por el artículo 332 del CPP.

[12] El artículo 332 del CPP ya citado también permite el uso del informe escrito con el propósito de demostrar o superar contradicciones.

desarrollo[13]. Antes del juicio, el acceso al informe escrito permitirá también a las partes decidir si es necesario encargar otros peritajes para que cumplan con un rol de refutación del que han examinado desarrollando una estrategia de confrontación más compleja que simplemente contra examinar (Horvitz y López, 2004, p. 298) y, en todo caso, permitirá controlar en cualquier momento el proceso de razonamiento del perito y verificar si se ha sometido a las reglas de la ciencia, arte u oficio que desempeña para poner en evidencia debilidades y problemas en cualquier etapa en que dicho informe se intente utilizar (Aguilera y Galleguillos, 2025, p. 64; Cerda, 2010, p. 457; Horvitz y López, 2004, p. 297).

Como se puede concluir, las diversas contribuciones del informe lo transforman en un instrumento clave para el funcionamiento del sistema. Por lo mismo, no cualquier formato o contenido permite que cumpla con su rol de favorecer un control efectivo de la prueba pericial. Por otra parte, aun cuando el contenido sea adecuado, el momento u oportunidad de acceso al mismo será una dimensión relevante para asegurar que esa función sea cumplida.

2.2. Estándares, propuestas y regulaciones de derecho extranjero en materia de informes periciales escritos

La comprensión del rol clave que cumplen los informes periciales escritos y la mayor conciencia sobre los riesgos en el uso de esta prueba, ha generado el desarrollo de diversos estándares, propuestas y nuevas regulaciones legales en el derecho comparado que han precisado sus contenidos mínimos y oportunidad de producción. El objetivo de esta sección es revisar algunos de ellos a modo ejemplar para observar el tipo de exigencias que hoy se discuten debiera satisfacer un informe escrito para cumplir con su rol. Esto permitirá luego contrastar estos desarrollos con las reglas y la práctica chilena. Se dividirá el análisis en tres aspectos: contenidos del informe (incluyendo documentación y registro en su elaboración); información sobre la idoneidad y antecedentes de quien realiza el informe; oportunidad y acceso.

2.2.1. Contenidos básicos del informe y exigencias documentación y registro en su elaboración

En su famoso informe sobre el estado de las ciencias forenses en los Estados Unidos, la Academia Nacional de Ciencia (2009) (en adelante NCR-NAS) de dicho país identificó la necesidad de elevar los estándares de reporte de los resultados de investigación pericial al constatar deficiencias en la materia[14]. Refiriéndose a los informes

[13] Por ejemplo, la Corte Suprema en Chile se ha pronunciado proscribiendo aportes de información del perito que no estén debidamente incorporados en su informe pericial (Rol n° 2866-2013) (Fierro y Walker, 2024, pp. 358-359).

[14] Ese hallazgo no parece ser único de los Estados Unidos. Para el caso de Australia (Edmond 2013, p. 11).

de laboratorio producto de análisis científico, la NCR-NAS (2009) estableció como un estándar general la necesidad que los informes fueran «completos y minuciosos» (p.186). Agregó que, como mínimo, ellos debieran describir «…los métodos y materiales, procedimientos, resultados y conclusiones...» (p. 186) Además, debieran identificar cuando correspondiere «…las fuentes de incertidumbre en sus procedimientos y conclusiones junto con las estimaciones de su escala (para indicar el nivel de confianza en los resultados)» (p. 186). El reporte de la NCR-NAS (2009) asume que el informe pericial no puede contener el mismo nivel de detalle que es esperable de un artículo académico, pero establece que «…debe tener un contenido suficiente que permita a un lector no científico comprender lo que se ha hecho y permitir un escrutinio informado y sin sesgos sobre sus conclusiones» (p. 186).

Como se puede observar, buena parte de las exigencias planteadas por el informe NCR-NAS (2009) se orientan directamente a mejorar las condiciones de control de la prueba pericial. Si bien están formuladas como estándares generales que requieren de mayor precisión, apuntan a que los contenidos de los informes sean exigentes y cubran diversos aspectos que permitan tener una idea clara de cómo se llegó a los resultados o conclusiones que presenta y sus fundamentos.

A partir de esta recomendación, la NCFS (2015) elaboró un documento destinado a darle mayor precisión a los alcances de los estándares propuestas por la NCR-NAS (2009). En este documento se realizan recomendaciones para los proveedores de servicios forenses y de medicina forense de los Estados Unidos sobre documentación, registro y contenidos de los informes periciales.

Un aporte muy significativo del documento de la NCFS (2015) es que distingue tres niveles de información que deben ser considerados en su conjunto en la elaboración de los informes periciales. En uno primero, que denomina documentación, se establece la necesidad de desarrollar políticas escritas para registrar todos los pasos y operaciones llevadas adelante en la construcción del informe. Sin ese registro, se puede perder información clave de fundamento del informe y que luego permita su control. La documentación es la base de los dos niveles posteriores.

En el segundo nivel -archivo, registro o expediente del caso- se establece que quien elabora el informe debe mantener un archivo con todos los documentos específicos del caso o examen generados en su desarrollo, incluyendo registros administrativos, resultados de exámenes y todo otro registro técnico o comunicaciones (electrónicas o físicas) generadas o recibidas por el o la experto que pertenezcan al caso concreto[15]. La idea de diferenciar el informe y el archivo del caso se presenta frente a la dificultad práctica que un dictamen escrito pueda contener la potencial gran cantidad de material que podría producirse/utilizarse en la elaboración de un reporte pericial. La

[15] La NCFS (2015) también establece que documentos de tipo genérico, es decir, no referidos al caso específico, por ejemplo, los estándares de procedimientos operativos o las definiciones básicas de los conceptos utilizados, podrían también ser parte del archivo del caso o, al menos, estar a disposición y ser de fácil acceso, por ejemplo, al estar en una página web o poder accederse a ellos previa solicitud.

NCFS especifica que el estándar a satisfacer del archivo o expediente del caso es que este debiera:

> …contener toda la documentación que permita a otro analista o científico, que cuente con el entrenamiento y experiencia adecuada, comprender y evaluar todo el trabajo realizado e interpretar y analizar en forma independiente los datos disponibles y extraer conclusiones para revisión de pares que han sido por largo tiempo un pilar de la ciencia y el método científico (NCFS, 2015).

Por cierto, también se manifiesta preocupación para que el acceso al archivo del caso sea oportuno [16]. Debe notarse que se fija un estándar aún superior al del NRC-NAS que ponía acento en que un lector no científico pudiera comprender el informe y favorecer su escrutinio. Aquí, en cambio, se plantea un acceso a prácticamente todo el material vinculado a la elaboración de esta prueba que permita a un experto hacer una revisión de pares.

Finalmente, en un tercer nivel, se especifican contenidos mínimos que debiera tener el reporte o informe escrito. Si bien la NCFS (2015) reconoce que ellos, por su naturaleza, entregarán menos información que el archivo completo del caso, fija como estándar general que deben transmitir con precisión y claridad suficiente la información necesaria para servir a la ciencia y al sistema de justicia penal. Además, señala que ellos deben advertir en forma explícita que se complementan con el archivo del caso en donde es posible revisar todos sus registros específicos [17]. Desarrollando el estándar general, la NCFS especifica que los reportes o informes escritos deben contener como mínimo:

> …una declaración de la finalidad del examen, las pruebas y la interpretación de las pruebas; el método y los materiales utilizados; un resumen o una descripción de los datos o resultados obtenidos; las conclusiones o interpretaciones derivadas de los datos o resultados; los resultados, interpretaciones o conclusiones discordantes; y, cuando sea necesario para la interpretación de los resultados de las pruebas, las fuentes de incertidumbre en el procedimiento y las conclusiones, junto con estimaciones de su magnitud (NCFS, 2015).

Como se puede apreciar, diferenciar esta materia en tres niveles (documentación, archivo y contenidos), es clave para asegurar la existencia de información que en su conjunto es lo que permiten a un tercero, sea o no experto, disponer de elementos para evaluar y hacer un escrutinio del trabajo del experto.

En Chile y, en general en América Latina, el foco central parece haber estado en el segundo nivel y ello podría explicar algunas carencias de los informes escritos. En la contracara, varios de estos niveles comienzan a ser recogidos en propuestas o legislaciones comparadas, las que muestran regulaciones mucho más completas que la nacional.

---

[16] La NCFS (2016) desarrolló también un documento con recomendaciones al fiscal general de los Estados Unidos en materia de descubrimiento que se hace cargo de esta materia y sobre la cual se volverá un poco más adelante.

[17] El texto de la NCFS (2015) sugiere incluir una cláusula que señale: «Este informe no contiene toda la información necesaria para evaluar independientemente el trabajo o interpretar en forma independiente los datos. Esta evaluación requiere la revisión del archivo del caso».

Un ejemplo paradigmático de derecho extranjero que vale la pena revisar en esta materia está constituido por la regla 19.4 de las Criminal Procedure Rules de 2020 de Inglaterra y Gales[18]. Esta disposición especifica los contenidos básicos del informe pericial recogiendo y expandiendo las propuestas que hemos visto desarrolló el NRC-NAS (2009) y la NCFS (2015) de los Estados Unidos. Dejando de lado por un momento los contenidos vinculados a los antecedentes del experto, la regla contiene ocho exigencias básicas (letras b, c, d, e, f, g, h, i), algunas de las cuáles se dividen en varios requerimientos específicos a la vez (letras e, f).

Sin entrar al detalle de cada una, destacan especialmente tres tipos de exigencias que se formulan con bastante especificidad en distintas letras de la Regla 19.4. Una primera pone fuerte énfasis en la identificación precisa de las fuentes de información utilizadas o en las que se basa el informe, ya sea de bibliografía u otra naturaleza (letra b), los hechos que le han sido comunicados al experto para su elaboración (letra c), los hechos que son de su propio conocimiento (letra d), y, la identificación cuando sus opiniones o inferencias han sido basadas en una representación de hecho u opinión realizada por otra persona (letra e). Sin información precisa en estos aspectos resulta difícil o casi imposible para un lector externo controlar adecuadamente a esta prueba. Por ejemplo, sin ella no es posible identificar potenciales sesgos cognitivos del perito -como el sesgo de información contextual- que podría incidir decisivamente en la calidad y confiabilidad de su opinión o conclusión[19].

Un segundo grupo de exigencias se asocia a la idea de reflejar con transparencia en el informe la existencia de cualificaciones o matices a su opinión y/o conclusión, ya sea a través de la identificación de opiniones diferentes sobre el punto (letra f) o de la existencia de reservas en sus opiniones (letra g). Se trata, nuevamente, de un aspecto clave para identificar potenciales debilidades de un informe o para explorar explicaciones alternativas a las que éste llega (Edmond, 2013, pp. 3-4).

---

[18] *https://www.legislation.gov.uk/uksi/2020/759/rule/19.4* Esta regla se complementa con varios otros instrumentos y recomendaciones, entre ellos, con las orientaciones desarrolladas por el Forensic Science Regulator (especialmente su número 4 actualizada en julio de 2024: *https://www.gov.uk/government/publications/expert-report-content-issue-4/expert-report-content-issue-4-accessible*) y las guías desarrolladas por el Crown Prosecution Service (CPS) de Inglaterra y Gales, especialmente la destinada al descubrimiento probatorio (en: *https://www.cps.gov.uk/legal-guidance/expert-guidance-disclosure-unused-material-and-case-management*).

[19] La evidencia muestra que la exposición de los expertos a información del caso que es irrelevante para su tarea puede afectar la confiabilidad y precisión de sus opiniones y conclusiones (NCR-NAS, 2009; NCFS, 2015; Dror, 2020; Edmond *et al.*, 2014; Vedrevelt *et al.*, 2024, p. 712) y también que las disciplinas que requieren considerar información contextual para construir sus informes son de aquellas en donde se documentan mayores problemas potenciales de sesgos cognitivos. (Morgan, 2023b, pp. 22-23). Por otra parte, la evidencia muestra que no solo el tipo de información al que acceden los peritos puede generar este efecto, sino que también el orden o secuencia en que dicha información es presentada juega un rol crítico (Dror, 2021, p. 1; Edmond *et al.*, 2014, pp. 9-13; Quagley-Macbride *et al.*, 2022).

Finalmente, es necesario prestar atención a la exigencia general de la letra h) que requiere al informe contener «…información que el tribunal pueda necesitar para decidir si el dictamen del perito es suficientemente fiable para ser admitido como prueba». Esta regla que, en principio parece algo abstracta y general, es la que permite construir las exigencias de especificación de los procedimientos y métodos utilizados en la elaboración del informe. En efecto, esta regla es complementada por las Criminal Practice Directions[20], las que dedican una porción bastante extensa del desarrollo destinado a la prueba pericial a los criterios o factores que deben considerar las cortes para decidir admitir una prueba de esta naturaleza. La directiva 7.1 desarrolla estos aspectos. Así, respecto a la metodología se incluyen cuestiones tales como la validez de la metodología empleada por el experto (7.1.2.b), el grado de precisión o el margen de incertidumbre que afecta a la exactitud o fiabilidad de los resultados (7.1.2.d), si los métodos usados por el perito se han ajustado a la práctica establecida en la materia (7.1.2.i), entre varias otras. Para que estos aspectos puedan ser evaluados por un tribunal, se debe tratar de información disponible, en principio, en el informe escrito[21].

A nivel de propuestas, se revisa brevemente la elaborada en forma reciente por Vázquez (2022) para México por su relevancia regional. Ella considera varios ejemplos de regulaciones y propuestas de derecho extranjero, incluyendo textos de América Latina como el artículo 226 del Código General del Proceso de Colombia de 2012 y el artículo 473 del Código de Proceso Civil de Brasil de 2015. A su vez, revisa el Anteproyecto de Ley de Enjuiciamiento Criminal de España de 2020 y regulaciones de Reino Unido y Australia (Vázquez, 2022, pp. 17-25).

Vázquez (2022) estima que, para que un informe pericial permita fundar decisiones judiciales compatibles con el debido proceso, debe incluir al menos seis aspectos de contenido básico. Su propuesta integra con bastante fidelidad las exigencias que se han planteado en los ejemplos revisadas en su trabajo y en los estándares y ejemplos de derecho extranjero analizados a modo ejemplar en esta investigación. Dos de ellas se refieren a exigencias de información vinculadas a las calificaciones o idoneidad de los expertos (2 y 3) que quedan fuera del análisis por el momento. Las otras cuatro incluyen contenidos del informe propiamente tal. Estos son resumidamente: (1) la descripción del objeto del peritaje y la información relevante que el perito ha recibido sobre el caso, con determinación de la hipótesis que plantea corroborar; (4) la

[20] Las directivas son reglas emanadas del Lord Chief Justice de Inglaterra y Gales y deben ser leídas en conjunto con las Reglas de Procedimiento Penal. Su última versión es del año 2023. *https://www.gov.uk/guidance/rules-and-practice-directions-2020* . Un análisis de las directivas en este punto, pero en una versión previa a la vigente puede verse en (Keane y McKeowon, 2020, pp. 665-668) y (Choo, 2009, pp. 337-340).

[21] El Forensic Science Regulator (en adelante FSR) de Reino Unido también elaboró una guía el 2020 sobre el contenido de los informes periciales. Junto con elementos similares a los ya descritos, plantea exigencias de información generales sobre la institución a la que pertenece el experto, por ejemplo, de si trata de una acreditada con estándares internacionales forenses como la norma ISO 17.025. Justifica esa exigencia en la cláusula que establece que el juez debe considerar toda evidencia que permita determinar la confiabilidad del peritaje para admitirlo (FSR, 2020, p. 22).

descripción clara y comprensible de los métodos, técnicas, teorías u otras empleadas para analizar el caso, señalando la mejor evidencia disponible sobre su validez y fiabilidad; (5) la referencia detallada a todo el conjunto y origen de los datos, así como los razonamientos efectuados durante su análisis (incluyendo documentación sobre todo aquello que se originó o empleó para el análisis del caso); (6) las conclusiones a las que llegó el perito basadas en las exigencias previas (p. 27 y pp. 28-61).

Según se puede observar, esta propuesta incluye exigencias importantes que parecen elevar el estándar de las regulaciones disponibles en la región y, como se verá, claramente en el caso de Chile. Especial énfasis nuevamente aparece en la identificación de la información utilizada, los métodos y procedimientos, así como el razonamiento y conclusiones del experto. Todos ellos suponen políticas claras de documentación previa en la línea del texto de la NCFS (2015).

La propuesta de Vázquez (2022) también incluye el incorporar un anexo al informe que contenga, en caso de haberse empleado, entrevistas, fotografías, estudios clínicos y demás información relevante obtenida sobre el caso en la elaboración del dictamen (p. 28). Este anexo pareciera ser el equivalente funcional del archivo del caso exigido por los estándares de la NCFS (2015) y comienza a consolidarse como exigencia compartida a nivel comparado [22].

2.2.2. Información relevante del experto

La posibilidad de hacer un escrutinio estricto de la prueba pericial exige también acceder a diversa información acerca de quien lo ha elaborado de manera de establecer si cuenta con las competencias requeridas por la ciencia, técnica o disciplina respectiva. Idealmente, también se debiera disponer de cualquier otra información adicional relevante sobre el experto que pudiera poner en juego la confianza en el peritaje o de su credibilidad o imparcialidad. Piénsese nuevamente en lo que ocurre con la presencia de sesgos cognitivos como, por ejemplo, los denominados sesgos motivacionales como el sesgo adversarial o de asociación, descrito como «La tendencia natural para hacer algo que sea útil para quien te emplea y remunera adecuadamente» (Paccioco, 2009, p. 577). Sin contar con información sobre este tipo de cuestiones que van más allá de las credenciales del experto se hace muy difícil sino imposible detectar su presencia y, por lo mismo, hacer un escrutinio intenso de esta prueba.

Recogiendo esta lógica, es posible identificar diversas propuestas y reglas que se hacen cargo de estas exigencias. Así, las recomendaciones sobre descubrimiento de evidencia forense de la NCFS (2016), siguiendo a la reglas federales de proceso civil de los Estados Unidos, establece que se debe entregar información sobre las califi-

[22] La exigencia de un anexo aparece mencionada en forma explícita en el artículo 482.2 del Anteproyecto de Ley de Enjuiciamiento Criminal de España de 2020 y en forma implícita en el artículo 226.10 del Código General del Proceso de Colombia, ambos analizados en el texto de Vázquez (2022).

caciones profesionales de los peritos, incluyendo la lista de las publicaciones que ha sido autor en los últimos diez años; una lista de todos los casos en los cuatro años previos en que el perito hubiera prestado declaración en juicio o en etapas previas; y, una declaración sobre la compensación recibida por el perito en la elaboración del informe (NCFS, 2016)[23]. Como se puede observar, se incluye información adicional a la referida a las tradicionales calificaciones que permite conocer opiniones previas del experto y antecedentes para evaluar potenciales sesgos a propósito de las condiciones de contratación.

Estas exigencias son profundizadas por las Criminal Procedure Rules de 2020 de Inglaterra y Gales y las directivas de 2023 que las complementan. Como ya se mencionó, la regla 19.4 letra a) establece el deber de incluir en el informe antecedentes sobre la calificación del perito, su experiencia relevante y su acreditación. Esta exigencia es complementada por un deber especial establecido en la regla 19.3(3)(c)[24], que solicita que la parte adjunte al dictamen cualquier información que tenga conocimiento y que pueda razonablemente considerarse susceptible de menoscabar la fiabilidad del dictamen pericial o la credibilidad o imparcialidad del perito. La directiva 7.1.4 complementa a lo anterior estableciendo que debe revelarse toda la información potencialmente relevante, tanto en relación con el perito como en relación con cualquier empresa u otro organismo con el que trabaje el perito, como empleado o en cualquier otra calidad[25]. Luego entrega a modo ejemplar un listado de doce aspectos que podrían tener relevancia, incluyendo: información sobre honorarios acordados; sobre potenciales conflictos de interés; noticia sobre comentarios judiciales adversos recibidos por el perito en otros casos; anulaciones de casos basadas en deficiencias en el testimonio del perito; procedimientos disciplinarios o críticas recibidas por sociedades o autoridades profesionales; condenas penales que puedan afectar la honestidad e integridad profesional del experto; historial de fracasos o malos resultados en evaluaciones de calidad o competencia del perito; entre otros.

Como se puede apreciar, estas reglas ponen un fuerte énfasis en contar con información mucho más rica que las tradicionales exigencias de calificación profesional como la identificación de los estudios, títulos profesionales y especialidades. Esto surge por dos razones que se entrelazan. Por una parte, la existencia de mayor consciencia sobre el impacto que tienen los sesgos cognitivos en la calidad y confiabilidad

[23] Interesante constatar que el artículo 226 del Código General del Proceso de Colombia, con seguridad por influencia de las leyes federales de proceso civil de los Estados Unidos, plantea exigencias muy similares en sus numerales 4 y 5. Agregando en su numeral 6 información si ha sido contratado en procesos previos o en curso por la misma parte o apoderado, información relevante para controlar potenciales sesgos de asociación mencionados.

[24] *https://www.legislation.gov.uk/uksi/2020/759/rule/19.3*

[25] Debe recordarse, como ya fue mencionado, que las orientaciones del FSR (FSR, 2020) sobre este punto también establecen también deberes de información referidos a la institución a la que pertenece el experto, por ejemplo, de acreditación con estándares internacional, en sus orientaciones 14.2.4 y 14.2.5.

de la prueba pericial y, por el otro, el que las propias reglas de procedimiento definan [Regla 19.2 (1)] que el deber central de los expertos es ayudar a los tribunales emitiendo informes objetivos y sin sesgos y dentro del ámbito de su especialización (The Law Comission, 2011). Esta mayor consciencia, expresada también normativamente en deberes para los expertos, se traduce luego en exigencias más altas de descubrimiento que permitan efectivamente controlar la calidad de esta prueba en esta dimensión[26].

La propuesta de Vázquez (2022) en este punto es un poco menos ambiciosa y, en alguna medida, todavía apegada a las exigencias más tradicionales de calificación profesional/técnica. Así, en su numeral (2) propone que el dictamen contenga información que permita considerar que el perito es experto en el ámbito específico relevante del objeto del peritaje y en el uso de métodos o técnicas requeridas para su desarrollo. En el numeral (3) se extienden estos deberes a otras personas que hayan participado en el análisis pericial realizado (p. 27)[27]. Aun en contexto de un desarrollo más escueto en este punto, destaco la diferenciación que se plantea entre credenciales profesionales del experto (en el ámbito específico del informe) y en el manejo de los métodos o técnicas específicas aplicadas al caso. Como muestra la literatura más reciente, la especialidad de los expertos suele ser en ámbitos muy específicos de conocimiento (Martire *et al.*, 2025) por lo que acreditaciones generales de área de conocimiento normalmente debieran ser consideradas insuficientes como prueba de su idoneidad.

2.2.3. Oportunidad y acceso al informe y sus archivos o registros

El control de la prueba pericial supone contar también con reglas que aseguren el acceso a la información de manera oportuna. En esta dimensión es difícil encontrar un estándar o regla única debido a las enormes diferencias de diseño procesal existentes en el ámbito comparado. Más allá de ellas, un régimen que pretenda generar un entorno de control adecuado de la prueba pericial debe resolver esto en forma consistente a las dimensiones previas.

Por ejemplo, en las recomendaciones de la NCFS (2016) se señala que la entrega del informe a las partes y al juez debe realizarse en un tiempo razonable antes del juicio, entendiendo que ello significa al menos varias semanas antes de dicha audiencia

[26] Esta lógica también fue reconocida por el informe de la Law Commsission (2011) en el que se planteó que, en la medida que las exigencias de admisibilidad de la prueba pericial se hicieran más estrictas, eso suponía mayores niveles de descubrimiento para que las partes y los jueces puedan adoptar decisiones correctas (párr. 7.35, p. 116).

[27] En el análisis específico de estos puntos de la propuesta menciona la exigencia de dar a conocer las publicaciones de los últimos años en el área relevante para el peritaje, pero no aborda elementos adicionales vinculados a los conflictos de interés o problemas que puedan afectar la objetividad o credibilidad del perito (Vázquez, 2022, pp. 33-37).

y con tiempo suficiente para que la defensa consulte y/o contrate asistencia experta (p. 2). Adicionalmente, la recomendación a los fiscales es también a asegurar el acceso de la defensa al archivo del caso, de manera de permitir revisar las fuentes de información que le dieron sustento al informe y que no están contenidas en forma completa en su desarrollo (NCFS, 2016, p. 2). Finalmente, se recomienda al fiscal general autorizar a los fiscales federales condicionar descubrimientos adicionales a la defensa a cambio que ella proporcione acceso similar respecto a la prueba pericial que pretende introducir (NCFS, 2016, p. 2). Esta sugerencia complementa a la regla 16(b)(1)(c) de las Federal Rule of Criminal Procedure que establecen deberes básicos de descubrimiento de la defensa, incluyendo las opiniones que el perito manifestará en juicio y los antecedentes sobre sus calificaciones profesionales, publicaciones y comparecencias previas[28]. Esta última recomendación muestra un punto de mucha relevancia. Si la necesidad de asegurar un mayor control de la prueba pericial surge de algunos riesgos que esta presenta debido a su particular naturaleza, las exigencias de contenido y oportunidad que hemos revisado debieran extenderse a toda prueba pericial que se intente utilizar y no sólo a la de los fiscales.

La regla 19.3(3) las Criminal Procedure Rules de 2020 de Inglaterra y Gales contempla un estándar similar al propuesto por la NCFS (2016) al establecer que quien quiera introducir prueba pericial a juicio debe entregar copia del informe a las partes y a los jueces tan pronto como sea posible. Llamo la atención a que esta regla es aplicable a todas las partes que en un caso penal quieran beneficiarse del uso de prueba pericial. Las directivas de 2023 complementan esto con la directiva 7.1.6 que establece que si el tribunal toma conocimiento que existe información sobre el perito o el peritaje que estaba dentro de aquellas que deben ser objeto de descubrimiento y no lo ha sido, podría potencialmente excluir esa prueba. Esto último agrega una dimensión relevante de análisis: las consecuencias procesales del incumplimiento de los estándares de descubrimiento. Por su relevancia, una infracción de estas debiera tener consecuencias de peso, como, por ejemplo, la potencial exclusión que plantea la directiva.

Para cerrar esta sección, los estándares, propuestas y regulaciones brevemente revisadas a modo ejemplar muestran cómo la admisión de prueba pericial en el proceso penal comienza a ser objeto de mayores exigencias en los niveles de información que deben estar disponibles en forma previa a su uso. Estos exceden a las tradicionales de acreditar de manera general la actividad profesional o técnica del experto y los contenidos, más o menos resumidos, solicitados al informe escrito.

[28] Texto completo en: *https://www.law.cornell.edu/rules/frcrmp/rule_16*

3. LA REGULACIÓN LEGAL DEL INFORME PERICIAL EN EL CPP EN CHILE

Estándares como las analizadas en el capítulo anterior no son desconocidos en Chile. Autores han planteado diversas propuestas de estructuración de los informes periciales que parecen reconocer estos desarrollos de derecho extranjero. Por ejemplo, se ha descrito al informe pericial como un documento «… cuyo objetivo es informar y demostrar de forma empírica las operaciones practicadas de manera que sea completo, comprensible y por sobre todo entendible para cualquier lector que no domina la ciencia, arte u oficio del perito que lo elaboró» (Cáceres, 2023, p. 18). Se pueden apreciar muchos elementos del estándar general planteado por la NRC-NAS (2009). Luego, el mismo autor, a partir de la revisión de disposiciones de legislaciones de América Latina y Europa, formula una propuesta detallada de estructura y contenidos del informe (Cáceres, 2023, pp. 87-99) e incluso una de reforma legal del artículo 315 del CPP que permita contener los mínimos que estima necesario debieran incluir (Cáceres, 2023, pp. 102-103). Un análisis de ambas permite reconocer a diversos elementos que ya han sido recogidos en el derecho extranjero.

Unos años antes, este tema también fue foco de preocupación y análisis en el trabajo de González-Wilhelm (2018). En él sostiene que el informe debe cumplir con un estándar básico dando «… cuenta paso a paso de manera narrativa y pormenorizada sobre cada una de las operaciones practicadas, los resultados obtenidos y las conclusiones que a partir de ello puedan ser formuladas» (González-Wilhelm, 2018, p. 301). Agrega, además, que para cumplir con este estándar

> No basta con hacer un mero listado de hallazgos ni suministrar una simple opinión u orientación. El perito deberá justificar siempre de manera sólida y coherente todo juicio que emita, apoyándolo en consideraciones respaldadas idealmente por bibliografía y argumentaciones racionalmente deducidas (González-Wilhelm, 2018, p. 301).

A partir de este estándar y de la revisión de diversos ejemplos de regulaciones de América Latina en materia penal, presenta una propuesta relativamente detallada de estructura y contenidos del informe escrito capaces de satisfacerlo (González-Wilhelm, 2018, pp. 310-331). Nuevamente, se puede identificar que está integrada por varios componentes recogidos en los ejemplos revisados. De hecho, en varios se cita bibliografía de países como los Estados Unidos e Inglaterra.

Estas exigencias también han sido recogidas en algunas regulaciones administrativas del Servicio Médico Legal (en adelante SML) que establecen protocolos para la realización de distintos tipos de peritajes. Así, por ejemplo, tratándose del protocolo de autopsia simple, se plantea que el informe debe «expresar los hallazgos, razonamientos y conclusiones en forma clara y precisa», además, debe «ser completo, detallado, legible, objetivo y comprensible no sólo para otros médicos, sino también para lectores sin información académica». El protocolo desarrolla también un listado de mínimos a contener en el reporte (SML, 2021).

Más allá de estas propuestas, la legislación chilena se hace cargo del informe pericial y su descubrimiento en tres regulaciones diversas. La exigencia de contenidos está contemplada en el artículo 315 del CPP. La información para acreditar idoneidad del experto que lo elaboró en el 314. Finalmente, el acceso al informe, en reglas generales de descubrimiento probatorio que incluyen, entre otras, a los artículos 182 y 260 del CPP. A continuación, se revisarán brevemente estos tres niveles, recurriendo tanto a la doctrina disponible como a la escasa jurisprudencia que sobre ellos hay a nivel de tribunales superiores[29].

3.1. Los contenidos del informe pericial escrito (artículo 315 del CPP)

La regla básica en materia de contenido de los informes periciales está regulada en el artículo 315 del CPP que establece ellos deben incluir: (a) la descripción de la persona o cosa que es objeto del peritaje; (b) la relación circunstanciada de todas las operaciones prácticas y su resultado; y, (c) las conclusiones que formularen.

El origen de esta disposición se encuentra en el artículo 237 del Código de Procedimiento Penal de 1906, que tiene una redacción idéntica en referencia a los contenidos exigidos a este instrumento[30]. Se trata de una fórmula de redacción que se repite en la región. Así, por ejemplo, el Código Procesal Penal Modelo para Iberoamérica de 1989 en su artículo 187 establece exigencias muy similares y lo mismo ocurre en varios códigos reformados de América Latina[31]. Lo que explica esta uniformidad es que todos comparten como fuente básica a la legislación procesal penal de España del siglo XIX. En efecto, los contenidos son muy similares también a los del artículo 478 de la Ley de Enjuiciamiento Criminal de 1882. Este texto, por su parte, tiene como

[29] En materia se realizó una revisión sistemática de sentencias de la sala penal de Corte Suprema del período 2015-2025 (mayo) a partir de la Base Jurisprudencial del Poder Judicial (*https://www.pjud.cl/portal-jurisprudencia*) de esta manera de contar con una visión actualizada del estado actual de la jurisprudencia en el país. Así, el 5 de mayo de 2025 se hizo un requerimiento a dicha base utilizando como filtros ¨Corte Suprema¨ y ¨Segunda sala (penal)¨, además del concepto del tesauro ¨Informe pericial¨. Se obtuvieron 116 resultados. Dicha búsqueda fue complementada entre los días 5 y 8 de mayo de 2025 con nuevos requerimientos a la base sin utilizar su tesauro sino con conceptos libres: ¨prueba pericial¨ ¨perito¨ y ¨peritaje¨. Se obtuvieron 34 sentencias adicionales. Así, se contó con una base de 150. Luego de la revisión de contenidos esa cifra bajó 11 fallos que tienen pertinencia específica para este trabajo. Este material es complementado por diversos fallos recopilados no sistemáticamente por el autor tanto de la Corte Suprema como de Cortes de Apelaciones a lo largo de años recurriendo a distintas fuentes y en contexto de distintas investigaciones.

[30] El anteproyecto de enero de 1995 regulaba esta materia en su artículo 243 con el mismo contenido, haciendo referencia explícita al artículo 237 del CPP de 1906 (Quezada, 1994, pp. 364-365).

[31] Dentro de ellas, por ejemplo, el artículo 218 inciso segundo del CPP de Costa Rica, 234 del CPP de Guatemala, 411 del CPP de Panamá, 212 del CPP de República Dominicana y 250 del CPP de Provincia de Buenos Aires. Un análisis con más detalle de 10 códigos de la región en Duce, 2013, pp. 148-140. Revisando Perú, Ecuador, Colombia, Argentina y México: (Cácerea, 2023, pp. 49-56). Con otros siete países (González-Wilhelm, 2018, pp. 307-310), quien concluye se trata de exigencias legales bastante uniformes.

antecedente la disposición del artículo 375 de la Ley Provisional de Enjuiciamiento Criminal de 1872, primera que regula a nivel legal sistemáticamente el tema y que establece requisitos al informe escrito en el ordenamiento jurídico español[32].

Resalto que la legislación chilena actual encuentra su base, sin modificaciones relevantes, en una regulación que tiene más de 150 años y que fue desarrollada en un contexto procesal completamente diverso como era el sistema inquisitivo mixto o reformado instalado por las leyes españolas de 1872 y 1882. Ese modelo descansaba en otra estructura de control de la prueba (centrada en el rol del juez de instrucción). Una consecuencia obvia es que esto daría cuenta de un desfase regulatorio en Chile. Además, en el momento del diseño de esta disposición, el rol de la prueba pericial en el funcionamiento del sistema de justicia penal no tenía la relevancia que tiene en la actualidad y tampoco existía el nivel de consciencia de los riesgos a que su uso inadecuado nos expone.

Se si revisan los contenidos requeridos por el artículo 315 del CPP, se puede advertir se trata de una regulación que, en una lectura estricta, podría obligar a incluir mucha de la información necesaria para permitir evaluar adecuadamente el peritaje. Por ejemplo, la letra a) podría ser leída como una exigencia de explicitar con el detalle de propuestas y legislaciones revisadas a modo ejemplar las fuentes de información sobre las que se construye el peritaje. La exigencia de descripción de la persona o del objeto apunta precisamente en esa dirección. En consecuencia, una interpretación extensiva o de carácter teleológico de la disposición podría llegar a una exigencia más estricta. Algo similar ocurre con la letra b) respecto a las exigencias vinculadas a los contenidos metodológicos mínimos a incluir que hemos revisado es frecuente en el ámbito comparado.

En la otra cara de la moneda, el listado de contenidos del artículo 315 del CPP está muy lejos de tener la precisión y completitud que contemplan legislaciones extranjeras o en las propuestas analizadas. Por lo mismo, admite interpretaciones que podrían entender que un informe escrito cumple sí incluye contenidos básicos o mínimos que están muy por lejos de las exigencias y necesidades de información que han sido revisadas.

El análisis de la escasa doctrina en Chile que ha profundizado en el análisis de esta disposición da cuenta que ambas versiones están en disputa. La exigencia de especificar los métodos aplicados en el informe escrito constituye un buen ejemplo. De acuerdo con Cáceres (2023), la actual redacción no compele «…al perito a informar la aplicación u utilización de algún método o de las herramientas utilizadas para la ejecución de las operaciones» (p. 47)[33]. Efectivamente, la literalidad del texto no permite derivar de manera clara esa exigencia. Con todo, González-Wilhelm (2018)

---

[32] Los contenidos de esta disposición en: *https://legishca.umh.es/1872/12/22/1872-12-22-ley-provisional-de-enjuiciamiento-criminal/*

[33] El autor sostiene que en este escenario los protocolos que desarrollen instituciones a cargo de ejecución de peritajes debieran hacerse cargo de este vacío. En la propuesta de reforma legislativa que hace el texto incluye específicamente una mención a la necesidad de agregar un nuevo inciso al artículo

sostiene que, a partir de los tres numerales del artículo 315, se genera una obligación al perito de «dar cuenta que se ha aplicado el método científico» lo que estima sería un requisito esencial para considerar que un informe posee el carácter de prueba pericial (pp. 309-310). Esta conclusión, en cambio, podría abonar la tesis que la explicitación de la metodología sí es parte esencial de los contenidos del informe para el artículo 315 del CPP[34].

Más allá de la disputa en este punto ejemplar, ambos autores proponen una estructura de informe escrito que complementa la exigencia legal. Ambos estiman que su propuesta está pensando en un desarrollo que consideran mínimo, pero que, a la vez, contempla elementos mucho más específicos y detallados que los del artículo 315 del CPP[35]. Por ejemplo, se propone una exigencia de explicitación de la metodología aplicada (Cáceres, 2023, pp. 92-94); del marco teórico usado (González-Wilhelm, 2018, pp. 320-322); de la bibliografía que le da sustento (Cáceres, 2023, p. 99; González-Wilhelm, 2028, p. 329); la inclusión de anexos o documentación que incluyan materiales relevantes para la comprensión del peritaje y su interpretación (Cáceres, 2023, p. 98; González-Wilhelm, 2028, p. 330); entre otros.

Analizando las otras dimensiones acerca de los contenidos planteadas por la NCFS (2015), es posible notar que en Chile no existen reglas en el CPP que establezcan exigencias de documentación del proceso de trabajo que permite construir el peritaje[36] y tampoco sobre el archivo o registro del caso. Por lo mismo, estas quedan entregadas a lo que se pueda regular a nivel de protocolos de las instituciones que proveen informes, con el natural menor peso normativo, en caso de que sea necesario extraer consecuencias frente a potenciales casos de incumplimiento como veremos.

3.2. Los antecedentes del experto (artículo 314 del CPP)

Sobre el acceso a antecedentes del experto, la legislación chilena presenta una particularidad. Esta consiste en que las exigencias de información vinculadas al experto son reguladas en forma autónoma al informe escrito y, por lo mismo, en principio no existe deber de incluirlas en su texto. Así, una disposición diferente al artículo 315

315 que dé cuenta de la «…técnica, metodología o método reconocido por la comunidad científica aplicado…» (Cáceres, 2023, p. 103).

[34] Esta discrepancia para reflejarse también los protocolos desarrollados por el SML. Así, en aquel referido a los peritajes de salud mental, se incluye un capítulo de metodología en la estructura mínima del informe (Resolución Exenta n° 1501 de 2021, SML en tanto que el que regula a las autopsias simples no se incluye (Resolución exenta n° 905 de 2021, SML).

[35] Cáceres (2023) habla que su propuesta incluye la de una estructura homogénea con requisitos obligatorios mínimos a partir de la cual se pueden incorporar más elementos (pp. 87-102). González-Wilhelm (2018) intenta plantear un modelo que evite ritualismos de antaño y entregue flexibilidad en la medida que cumplan con unos mínimos (pp. 310-331).

[36] Sobre documentación las normas del CPP en la materia regulan las actuaciones del Ministerio Público (artículo 227) y las policías (artículo 228) y están referidas a dejar constancia de las actuaciones de investigación que realizan.

del CPP impone el deber de entrega de información sobre el perito como un requisito básico de admisibilidad de esta prueba. Se trata del artículo 314 inciso tercero del CPP, el que establece que la parte que intente introducir esta prueba a juicio debe acompañar comprobantes que acrediten la idoneidad profesional del perito para su discusión en la audiencia de preparación de juicio oral.

Varias observaciones. La primera, es que se trata de una exigencia para cualquiera de los intervinientes en un proceso penal que intente valerse de prueba pericial a juicio. Desde esta perspectiva se trata de una disposición que favorece el control. La segunda es que ella se hace cargo de una porción menor de la información relevante a considerar sobre el experto ya que se focaliza sólo en los antecedentes de idoneidad (profesional o técnica), pero no incluye otros aspectos claves [37]. Por ejemplo, aquellos asociados a su participación previa en el sistema (por ejemplo, listado de comparecencias en otros casos), sus publicaciones en el área o las remuneraciones percibidas por la realización del peritaje. Tampoco se incluye la entrega de información relevante que pudiera comprometer la imparcialidad del perito o mostrar la existencia de conflictos de interés, entre otras. Tratándose incluso de expertos que trabajan en organismos auxiliares del Ministerio Público, ha primado la visión que el artículo 321 del CPP hace innecesario acreditar su idoneidad, bastando el que sean parte del organismo respectivo para entender que se cumple con ese requisito (Aguilera y Galleguillos, 2025, p. 36; Cáceres, 2023, p. 48).

Una última observación se refiere al alcance limitado de la disposición desde el punto de vista de las hipótesis en que es exigible. Esta solo operaría en los casos en que algún interviniente intenta introducir la prueba pericial a juicio. En efecto, la disposición establece de manera explícita que el deber de acompañar los antecedentes de idoneidad del experto está condicionado a los casos en que se intente introducir esta prueba a juicio para efectos de su discusión en la audiencia de preparación de juicio oral. De esta forma, para todos los casos en que esta evidencia se ocupe en decisiones preliminares, no existiría un acceso garantizado a la información sobre la idoneidad del experto. Como ya fue mencionado, muchas de estas decisiones pueden ser críticas para el desarrollo del proceso.

---

[37] Pareciera existir acuerdo en la doctrina y la jurisprudencia que la exigencia de acreditación del artículo 314 se ha transformado en una relativamente formal, bastando con la presentación del currículo y copias simples de los títulos de los expertos. Por ejemplo, Corte de Apelaciones de Santiago, Rol n° 2130-2009, 21 de octubre de 2009; Corte de Apelaciones de Arica, Rol n° 145-2010, de 7 de julio de 2010; Corte Suprema, Rol n° 11515-2014, de 21 de julio de 2014. También sobre este punto (Aguilera y Galleguillos, 2025, pp. 35-37). Un caso más reciente, la Corte Suprema rechaza un recurso de nulidad fundado en violaciones del debido proceso al habérsele excluido un perito a la defensa por falta de información mínima sobre su idoneidad. En el caso, el perito fue presentado como un ingeniero en sonido, sin embargo, en la audiencia de preparación no se detalla ni se acredita si sus estudios fueron universitarios o técnicos ni el título específico. El juez de garantía excluye por falta de acreditación de la idoneidad y la Corte Suprema considera que ello es una correcta aplicación de las leyes procesales. Seguramente de haberse entregado copia de los títulos la exclusión no hubiera prosperado. Corte Suprema, Rol n° 32.047-2024 de 18 de octubre de 2024, considerandos 2° y 16°.

3.3. Oportunidad en el acceso al informe escrito y los antecedentes del experto

Mejoras en el registro, documentación y contenidos de los informes periciales pueden tener un impacto limitado sino van de la mano de reglas y prácticas que aseguren acceso oportuno a la información (Aguilera y Galleguillos, 2025, p. 63). Sobre esta materia, Chile cuenta con una regulación que se hace cargo de algunas de las dimensiones del problema, pero no todas. Los artículos 182 y 260 del CPP establecen un estándar de descubrimiento de los peritajes de la fiscalía que parece adecuado. En primer lugar, permiten que el imputado y su defensa accedan de manera completa a ellos e incluso obtengan copias. En segundo término, ese acceso debe producirse durante el desarrollo de la etapa de investigación y, en todo caso, cuando esta haya sido cerrada[38].

Las normas que regulan el descubrimiento de los peritajes de los querellantes (artículo 261 letra c y 259 letra f del CPP) y la de defensa (artículo 263 letra c CPP) ponen especial énfasis en la individualización del perito y sus calidades, pero son menos explícitas en lo que pasa con el informe escrito[39]. La oportunidad del descubrimiento también presenta algunas diferencias. Así, la querellante debiera descubrir su prueba al menos 15 días antes de realizarse la audiencia de preparación del juicio (artículo 261 CPP) y la defensa, en cambio, podría hacerlo incluso en la misma audiencia (artículo 263 CPP). Menores niveles de exigencia de descubrimiento sobre el contenido y oportunidad de entrega de los informes escritos son problemáticas ya que, como ha sido señalado, las necesidades de controlar a la prueba pericial también son extensibles a aquella que presentan querellantes y defensas, debiendo existir reglas y prácticas que favorezcan un escrutinio serio en esos casos.

Finalmente, tampoco existe una regulación completamente clara en Chile acerca de las consecuencias que tiene el incumplimiento de los deberes de acceso a la información, al no existir un tratamiento expreso del punto (Aguilera y Galleguillos, 2025, pp. 55-56). El caso más evidente de incumplimiento se da cuando se intenta

---

[38] A pesar de la claridad de estas reglas ha existido debate en la jurisprudencia sobre el momento en que se deben acompañar los antecedentes que acrediten idoneidad. En una decisión de la Corte de Apelaciones de Antofagasta se sostuvo que se cumple el deber al momento de la audiencia de preparación de juicio oral y no antes. Corte de Apelaciones de Antofagasta, Rol n° 276-2010, de 3 de septiembre de 2009. En otro de la Corte de Apelaciones de Talca, se rechaza un recurso de nulidad de la causal del artículo 374 letra c (haberse impedido a la defensa el ejercicio de las facultades que la ley le otorga) en un caso en el que el peritaje del Ministerio Público fue solicitado en el curso de la investigación, pero fue recibido por la fiscalía una vez cerrada esta, en todo caso antes de desarrollarse la audiencia de preparación de juicio oral. Corte de Apelaciones de Talca, Rol n° 830-2023 de 18 de agosto de 2023. En la misma dirección Corte de Apelaciones de Coyhaique, Rol n° 42-2023 de 21 de abril de 2023.

[39] Se ha sostenido que el deber que impone estas reglas es el de individualizar a los medios de prueba, pero no acompañarla, de manera que es posible y común en la práctica que ello ocurra al inicio de la audiencia de preparación de juicio oral (González-Wilhelm, 2018, p. 333).

introducir una prueba pericial a juicio sin haber acompañado oportunamente el informe. Aguilera y Galleguillos (2025) sostienen que lo que correspondería es declarar inadmisible el peritaje (pp. 55-56). Postura que tiene soporte en alguna jurisprudencia (Corte de Apelaciones de Chillán rol n° 93-2007 de 8 de agosto de 2007), pero como ha sido citado hay varios fallos de tribunales superiores que han admitido el uso de prueba pericial en casos en que el informe escrito fue entregado por el MP con posterioridad al cierre de la investigación en la medida que hubiera sido solicitado antes de esta.

Los casos más complejos, probablemente la situación más frecuente, se producen cuando existe un informe escrito, pero con un desarrollo precario que no satisface las exigencias del artículo 315 del CPP[40]. En este punto, se ha reportado que prueba pericial ha sido incluso admitida cuando el informe está constituido por un correo electrónico enviado por el experto al fiscal con un formato de desarrollo mínimo (González-Wilhelm, 2018, pp. 330-331). Se ha señalado también que, si bien el juez de garantía podría no admitir un peritaje con un informe con un desarrollo insatisfactorio, ello sería discutible (Cáceres, 2023, p. 47). El problema central para resolver los casos de informes escritos precarios o incompletos está en la falta de un acuerdo más preciso acerca de cuál es el alcance de las exigencias actuales del artículo 315 del CPP. Si esa norma puede ser satisfecha con una lectura minimalista, ello favorecerá la admisión de informes de mala calidad. Por otra parte, en este punto se notan algunas confusiones relativas al carácter complejo de la prueba pericial, al ponerse acento que la prueba a valorar sería la declaración en juicio y que, por lo mismo, defectos del informe no debieran impedir su admisión[41].

---

[40] Un caso interesante de la Corte Suprema se da en un recurso de nulidad en el que la defensa alegó como causal de vulneración de sus derechos fundamentales el que se haya admitido a juicio una prueba pericial cuyo informe no había consignado de manera evidente todas las operaciones practicadas por el experto para construir su opinión, vulnerándose al artículo 315 letra b) del CPP. La Corte considera infringida esa regla e incluso que ese defecto afectaría la exigencia de la letra c) del mismo numeral. Sin embargo, se rechaza la nulidad por considerar que se trató de una infracción no sustancial. Corte Suprema Rol n° 4712-2017 de 28 de marzo de 2017.

[41] En esta línea (Cáceres, 2023, p. 48). En la vereda opuesta (González-Wilhelm, 2018, pp.342-343). Una sentencia de la Corte Suprema al rechazar un recurso de nulidad interpuesto por la defensa por vulneraciones del debido proceso muestra también muestra la tensión que se produce entre las exigencias del artículo 315 del CPP y la declaración en juicio del experto. La Corte no da lugar al recurso basado en que en la audiencia de juicio el perito habría agregado conclusiones no contenidas en su informe señalando que «…de otra forma, los peritos se verían restringidos nada más que a reproducir literalmente el contenido de sus informes escritos en la audiencia, sin poder ilustrar a las partes ni al tribunal sobre otros aspectos relevantes para la resolución del juicio que se derivan o se hallen dentro del marco o contexto de las actuaciones realizadas como parte de la pericia encargada, escenario en el cual no se justificaría su comparecencia personal y bastaría la mera presentación del informe escrito, lo que, huelga señalar, se contrapone con la esencia misma de la arquitectura del juicio oral, donde se intenta que la verdad surja del debate que ahí se desarrolle mediante la interacción de los medios de pruebas con los intervinientes frente a los ojos de los jueces y no de la mera repetición textual y mecánica de los antecedentes recopilados durante la investigación, lo que, de aceptarse, desde luego importaría un paso atrás hacia el sistema inquisitivo de enjuiciamiento». Corte Suprema Rol n° 12.646-2019 de 25 de junio de 2019 (considerado séptimo).

Para cerrar este capítulo, se puede concluir que la legislación chilena está lejos de ofrecer una regulación clara y precisa sobre los informes escritos y, en general, sobre los contenidos que deben disponerse para facilitar un control intenso de esta prueba. Por vía interpretativa sería posible una lectura que eleve los estándares de manera que los informes cuenten con información suficiente y de calidad para favorecer su control en la línea de los desarrollos experimentados en el ámbito comparado. Con todo, la actual redacción permite también una lectura que podría dificultar esta tarea. Por lo mismo, conocer cómo en la práctica son entendidas estas exigencias es clave para saber si se está en presencia de un problema interpretativo o frente a uno que haga necesario discutir algunas reformas legales más profundas.

4. LOS INFORMES PERICIALES ESCRITOS EN LA PRÁCTICA DEL SISTEMA DE JUSTICIA CRIMINAL CHILENO

Este apartado tiene por propósito entregar una imagen de cómo en la práctica del sistema se componen los informes periciales escritos. Para estos efectos, en primer lugar, se dará cuenta de la escasa evidencia disponible en la materia. En la segunda parte, se expondrán los resultados de la investigación empírica exploratoria llevada adelante para la elaboración de este trabajo.

4.1. La evidencia disponible

En años recientes se ha acumulado un conjunto de trabajos con evidencia empírica de distinto carácter que sugiere los informes escritos presentan problemas. Muy ilustrativa es la investigación elaborada por la DECS (2017) basada en encuestas a jueces, peritos y abogados. Como punto de partida, la encuesta muestra que la evaluación sobre la calidad del trabajo realizado por los peritos que hacen jueces y abogados se distancia de manera significativa que la que ellos mismos hacen de su labor. Así, jueces y abogados del área penal le entregan una nota 4,7 y 4,8 respectivamente (en escala 1 a 7), los peritos penales se asignan un 6,38 (DECS, 2017, p. 27). Si bien, la evaluación general no es negativa, la diferencia de apreciación parece ser muy significativa y mostraría una mirada mucho más crítica de actores del sistema distintos a los peritos en relación con la calidad de trabajo de ellos.

Otras respuestas también entregan pistas de las razones que podrían tener jueces y abogados para hacer una valoración inferior a la que hacen los propios peritos de la calidad de su trabajo. Así, los abogados penales califican con un 3,6 la fiabilidad y validez de la metodología empleada en los peritajes y con un 3,9 la correspondencia entre premisas de los informes y sus conclusiones. En ambos casos se trata de calificaciones reprobatorias. Al justificar sus afirmaciones, la mala calidad de los informes escritos aparece como una cuestión de relevancia (DECS, 2017, p. 29). La visión de los jueces resulta ser un poco más positiva con calificaciones de 4,5 y 4,8 respectiva-

mente, pero igualmente expresando la existencia de importantes déficits de calidad de los informes escritos (DECS, 2017, p. 30). Finalmente, frente a la pregunta sobre si han tenido malas experiencias con peritos, incluyendo explícitamente la existencia de reiterados incumplimientos en su deber de acompañar informes o la presentación de informes de mala calidad, un 59,75% de los abogados y un 54,9% de los jueces señala que los han tenido con frecuencia. A ello se suma un 12,62% y 8,82% respectivamente en que se afirman que esos problemas se han generado en alguna ocasión (DECS, 2017, p. 30).

Si bien la investigación de la DESC no estuvo centrada en los informes escritos, sus resultados sugieren que la evaluación mediocre que realizan jueces y abogados penales sobre el trabajo de los peritos estaría fundada, entre otras cuestiones, en problemas asociados a la calidad de los informes que preparan. Esto se reitera en un estudio evaluativo del sistema procesal penal acusatorio a más de diez años de su puesta en marcha publicado el año 2017. En él se plantea, a partir de entrevistas realizadas a operadores del sistema, una visión crítica sobre la calidad de los informes periciales escritos elaborados por el SML, especialmente derivado de su carácter escueto y al carecer de desarrollo o profundidad (Arellano, 2017, pp. 175-176).

Otra investigación empírica (basada en entrevistas a jueces, abogados, fiscales y peritos) llegó a conclusiones similares. Un primer defecto identificado fue la frecuente falta de explicitación del método utilizado por el experto en su informe escrito, con lo cual las conclusiones de estos solían quedar con una justificación débil (Duce, 2018, pp. 66-67). A ello se sumaban otros problemas de calidad. Uno es el carácter incompleto de los informes (Duce, 2018, p. 67). Otro es la falta habitual de material básico de complemento utilizado en la elaboración del peritaje (transcripción de entrevistas, resultados de pruebas, dibujos, etc.) (Duce, 2018, p. 68). Finalmente, se identifican problemas de redacción y claridad en los contenidos (Duce, 2018, p. 68). Como complemento de lo anterior, la investigación detectó una práctica en juicio oral consistente con el hallazgo de que los informes suelen tener problemas de completitud y desarrollo. Se trata de lo se denomina peritos que «hablan más de lo que escriben», es decir, que en la audiencia de juicio agregan aspectos relevantes de su testimonio que no están consignados en el informe escrito, tratándose de un fenómeno frecuente y que recae en temas de importancia del caso (Duce, 2018, pp. 80-81)[42]. Este último fenómeno sugeriría también que efectivamente existen confusiones a partir de la idea que la prueba pericial es la declaración en juicio, considerando que ello libera al informe de un desarrollo completo y claro.

[42] En esta investigación se da cuenta de diversos fallos de la Corte Suprema en los que ha declarado la nulidad del juicio al estimar esta situación afecta al derecho a defensa de los acusados al impedirles preparar adecuadamente la confrontación del experto. Corte Suprema Rol n° 2866-2012 de 17 de junio 2013 y Rol n° 448-2015 de 11 de marzo de 2015. Esta jurisprudencia se habría matizado en un fallo reciente de rechazó un recurso en el que, a pesar de producirse este fenómeno, la Corte considera que no constituyó una infracción sustancial o de la entidad suficiente para dar lugar a la nulidad. Corte Suprema Rol n° 12.646-2019 de 25 de junio de 2019.

La falta de explicitación clara de la metodología utilizada por el experto ya había sido detectada previamente por otra investigación de tipo exploratoria centrada en peritajes psicológicos y psiquiátricos en las que se revisaron solicitudes de peritajes de los fiscales y los informes escritos de respuesta (Quijada, 2011, pp. 477-499). También el problema de falta de fundamentación adecuada de las conclusiones (González-Wilhelm, 2018, pp. 322-323).

Como puede observarse, a pesar de ser un área de poco estudio, la evidencia disponible muestra una imagen consistente: los informes escritos con frecuencia no parecieran satisfacer en la práctica a los estándares de contenidos revisados en el ámbito comparado y que, en una lectura estricta, podrían entenderse como los exigidos en Chile.

4.2. Resultados de una investigación empírica exploratoria

En esta sección se dará cuenta de los resultados de la investigación empírica exploratoria utilizando al efecto 90 informes periciales escritos requeridos por el MP (n: 33) y la DPP (n: 57) en el contexto de tramitación de casos penales a su cargo[43].

La investigación se estructuró sobre la revisión en profundidad de cada informe y la construcción de una base de datos en la que se incluyó dos tipos de información. En primer lugar, varios campos fueron dedicados a obtener aspectos identificatorios básicos de los informes. Además, se incluyó la revisión de diez aspectos de contenido a partir de las exigencias más frecuentes que se realizan en el ámbito comprado según se tuvo oportunidad de revisar, como también por algunos aspectos identificados como deficitarios en la evidencia disponible en Chile. La tabla n° 1 contiene en listado de aspectos en los dos ámbitos que fueron revisados.

[43] Los informes del MP fueron obtenidos de carpetas de investigación a disposición del autor en el contexto de investigaciones previas en la materia. Los informes de la DPP fueron proporcionados por la institución en el segundo semestre de 2024 a solicitud del investigador, bajo el requerimiento que se tratara de informes sobre las materias más frecuentes, recientes y que tuvieran diversidad regional. El análisis pondrá foco en resultados agregados, pero en el caso que se citen ejemplos, la información será presentada de manera anonimizada. Los informes serán identificados con un código genérico: los del MP como MP n° X y los de la DPP como DPP n° X.

Tabla 1 Elementos de los peritajes analizados

Aspectos identificatorios básicos	Exigencias más frecuentes en el ámbito comparado
— Asignación de n°	— Incluye descripción de la metodología usada en la elaboración del informe
— Objeto o materia del informe	— Incluye anexo con información relevante
— Fecha del informe	— Contiene la explicitación de la información revisada por el experto al elaborar el informe
— Ciudad de la firma del informe	— Hay referencia a las instrucciones del encargo de parte del mandante
	— Se explicitan los objetivos del informe
	— Hay conclusiones que guarden coherencia con el objetivo del informe y el encargo
	— Existe referencia a limitaciones (por ejemplo, márgenes de error) de la opinión elaborada
	— Hay presencia de datos curriculares del experto (generales o específicos)
	— Se utilizan citas/referencias y fuentes bibliográficas de respaldo
	— Extensión en número de páginas

La estrategia de recopilación de la información se realizó en un formato binario, es decir, si el informe incluye o no el ítem respectivo independiente de la calidad o profundidad que tenga en caso afirmativo. De esta forma se ha intentado evitar sesgos o subjetividades en la evaluación de los contenidos del informe, pero con un impacto en la calidad de la información recogida. Como se verá, solo en materia de información curricular se hizo una distinción en calidad de contenido al categorizar los casos que sí contaban con información curricular como detallado o genérico según se explicará. En materia de citas y referencias también se hizo el ejercicio de identificar el número de referencias y si ellas eran utilizadas en el desarrollo del texto o sólo incluidas como referencias al final de este.

El método utilizado tiene limitaciones de representatividad y en la profundidad y extensión de las conclusiones que son posibles de ser obtenidas. Por otra parte, pueden existir sesgos en la selección de informes enviados por el DPP como en la elección de los del MP. Con todo, se trata de una aproximación empírica exploratoria que no está disponible en el país y que, junto con las otras fuentes revisadas, permite entregar una primera imagen con plausibilidad empírica acerca de la práctica chilena en la confección de los informes escritos en materia penal.

En primer lugar, se presentarán algunos datos de caracterización de la muestra revisada y luego se abordarán los principales hallazgos respecto a la estructura y contenidos de estos informes.

4.2.1. Caracterización general de los informes revisados

La tabla 2 muestra el número y porcentaje de informes por especialidad o materia, siguiendo para estos efectos la clasificación que utiliza el SML para reportar los peritajes que realiza. Se ordenan, además, por la institución que los requirió.

Tabla 2 Número y porcentaje de informes analizados por materia e institución

Tipo	MP n° y (%)	DPP n° y (%)
Salud mental	5 (15,15%)	31(53,37%)
Tanatológicos	3 (9,09%)	4 (7,01%)
Clínicos	4 (12,12%)	12 (21,05%)
Laboratorios	8 (24,24%)	6 (10,52%)
Otros	13 (39,39%)	4 (7,01%)
Total	33 (100%)	57 (100%)

Fuente: elaboración propia

En el área salud mental, el total de los informes solicitados por el MP correspondió a peritajes de psicología forense. En los de la DPP, 18 (31,57%) correspondieron a esa materia y otros 13 (22,8%) a psiquiatría forense. Tratándose de la categoría otros solicitados por los fiscales, ellos correspondieron a informes: planimétricos (cuatro, 12,12%); balísticos (tres, 9,09%), documentales y de electro ingeniería (dos cada uno, 6,06%) y dactiloscópico y de tránsito (uno cada uno, 3,03%). En la DPP hubo un informe balístico, otro dactiloscópico, uno de tránsito y, finalmente, uno criminalístico (1,75% cada uno). Los informes provinieron de 30 expertos tratándose de los requeridos por el MP y 34 en el caso de los de la DPP[44].

La diferencia en el tipo de informes periciales utilizados por cada institución refleja de manera gruesa sus diversas funciones y cómo ellas inciden en sus necesidades probatorias. Los informes del MP muestran una distribución en varias categorías lo que se asocia a la carga de la prueba que tienen, allí donde los de la DPP presentan una fuerte concentración en los de salud mental que suele ser una materia relevante para la construcción de teorías del caso defensivas[45].

[44] En el caso del MP hubo tres peritos que se repitieron confeccionando dos informes cada uno. En la DPP hubo 12 expertos que hicieron más de un informe: seis realizaron dos peritajes, tres otros tres, dos confeccionaron cuatro y uno cinco. En relación con el sexo, 15 expertas que elaboraron informes a solicitud del MP fueron mujeres y 15 hombres. Tratándose de los peritajes de la DPP se trató de 14 mujeres y 20 hombres.

[45] Los datos de los peritajes encargados y pagados por la DPP en el año 2019 muestran que efectivamente los informes en salud mental representan un porcentaje importante de los que solicitan (25,8%), siendo la segunda categoría luego de los informes requeridos a asistentes sociales (58,8%) (Defensoría Penal Pública, 2020, p. 69).

Como elementos adicionales de caracterización se puede señalar que un 24,24% de los peritajes encargados por el MP lo fueron en regiones y el porcentaje restante en la Región Metropolitana de Santiago. Tratándose de los de la DPP un 63,15% en cambio fue firmado en una ciudad regional. Finalmente, un 66,66% de los peritajes del MP y un 98,75% de la DPP son del año 2020 o posteriores.

4.2.2. Principales hallazgos en materia de contenidos

Para efectos de facilitar la exposición y análisis se agrupan los hallazgos en dos temas: heterogeneidad en la estructura y contenidos de los informes, y falta de completitud en componentes relevantes.

4.2.2.1. Heterogeneidad de los informes

Un primer hallazgo es que se constata bastante heterogeneidad en la estructura y contenidos de los informes. En abstracto, esto podría ser un fenómeno esperable y en algún sentido consecuencia natural de la diversidad de disciplinas que proveen conocimiento experto al sistema de justicia penal. Por la naturaleza de su objeto de estudio, no debieran ser necesariamente similares informes de autopsia, de evaluaciones psicológicas o resultados de un laboratorio. En menor medida, la heterogeneidad también podría ser esperable en reportes de una misma disciplina y objeto, pero encargado por actores diferentes (MP versus DPP), reflejando así necesidades probatorias y de rol.

La investigación encontró una situación más problemática que la que sería esperable. En efecto, se pudo identificar diferencias sustanciales de estructura y contenidos incluso en informes de una misma disciplina y objeto encargados por el mismo interviniente. En algunos casos esas diferencias muestran incumplimiento de mínimos esperables de estos instrumentos. Eso sugeriría que las exigencias del artículo 315 del CPP no serían en la práctica entendidas de manera estricta y que la forma en que se interpretan permitiría la elaboración de textos con niveles muy disímiles de información, eventualmente con menos de la requerida o al menos esperable.

Un ejemplo de esta heterogeneidad se da en la comparación de tres informes de autopsia solicitados por el MP a la misma institución pública y producidos entre octubre de 2019 y marzo de 2021 por tres especialistas diversos[46]. En solo dos de ellos se mencionan los antecedentes revisados por el perito antes de realizar el informe. En uno solo se especifican las operaciones practicadas para efectuar el informe y en los otros esto no menciona o solo podría deducirse a partir del resto de la información. Solo uno contiene alguna referencia bibliográfica. Como se puede ver, existe diversidad importante de contenido. Con todo, en el ítem que se detecta mayor diferencia

[46] MP n° 1 (de 16 de octubre de 2019), MP n° 6 (de 27 de enero de 2021) y MP n° 18 (de 2 de marzo de 2021).

es en el formato y detalle en el que se presentan los resultados y las conclusiones de cada informe. Así, uno de estos se hace en forma detallada y narrativa de los mismos y en los otros, en cambio, son presentados de manera muy genérica, como un «mero listado de hallazgos»[47]. Todo esto se refleja en diferencias significativas de extensión. El más detallado tiene siete páginas y los otros tres y cuatro respectivamente.

Otro ejemplo interesante se obtiene de la comparación de cinco evaluaciones psiquiátricas encargadas por la DPP entre 2023 y 2024 a expertos diversos[48]. Un indicador de estas diferencias se expresa comparando su extensión. Así, el informe más breve tiene cinco páginas y el más extenso casi cuadruplica esa cantidad llegando a 19 (los otros tienen 14, 12 y 8 páginas). Detrás de esto existen niveles de desarrollo muy diversos en materias como: la explicación de la metodología; la exposición de los antecedentes relevantes del caso y en el diagnóstico; y, conclusiones que se formulan. Por ejemplo, en algunos casos la metodología, diagnóstico y conclusiones se expresan como un listado de afirmaciones sumamente breves (en estilo bullets) y en otros corresponde a un estilo narrativo. En la presentación de antecedentes relevantes, hay elementos que se diferencian entre los informes: tres incluyen transcripciones literales entre comillas acotadas de la entrevista con la persona objeto de la pericia, uno hace transcripciones extensas sin uso de comillas y un informe no contiene transcripciones de ningún tipo. Un solo informe incluye, por su parte, referencias bibliográficas.

Saliendo del análisis de estos casos ejemplares, el uso y forma de presentación de referencias bibliográficas constituye un buen indicador que muestra diferencias importantes en los contenidos en la globalidad de informes revisados. La necesidad que ellos identifiquen bibliografía o fuentes de apoyo ha sido enfatizada por la doctrina nacional como un componente necesario[49] y suele incluirse en disposiciones de derecho extranjero como la regla 19.4.b de las Reglas de Procedimiento Penal de 2020 de Inglaterra y Gales.

La investigación mostró un panorama lejano a esto, con diferencias muy significativas entre los informes solicitados por el MP y la DPP. En los del MP, solo en 9% (3) existe uso de fuentes de referencia y en los de la DPP dicha cifra se eleva a un 54,3% (31).

La forma en que se presenta la bibliografía también difiere de manera importante. En los informes del MP, en dos casos la referencia se utiliza como notas a pie de pá-

[47] Terminología utilizada por González-Wilhelm (2018) para describir críticamente la práctica de informes en el área médico legal (p. 301). Recordemos que sobre este punto el protocolo del SML exige al informe presentación detallada y comprensible de hallazgos, razonamientos y conclusiones (SML, 2021a).

[48] DPP n° 5 (abril 2023); DPP n° 6 (mayo 2024); DPP n° 13 (mayo 2024); DPP n° 14 (mayo 2024); y, DPP n° 30 (agosto 2023).

[49] La inclusión de bibliografía de apoyo en la construcción del informe sería un contenido obligatorio del mismo de conformidad a lo que postula (González-Wilhelm, 2018, pp. 329-330). Haciendo una crítica a partir de un caso específico en el que se detectan problemas de información y acceso a las referencias que podrían afectar el derecho a la defensa (Urrea, 2024, p. 276).

gina (cada uno con una cita e identificando a una fuente) y en uno se presenta como un listado de bibliografía al finalizar el texto. A todas luces se trata de un componente menor en este tipo de peritajes.

Los informes de la DPP muestran una realidad distinta, pero que no deja de ser problemática. Se pudo detectar tres estilos muy diversos de presentación. En doce informes las referencias bibliográficas aparecen como notas a pie de página exclusivamente. En otros seis, además de la nota, se incluye la fuente al final del informe en un listado bibliográfico. Finalmente, en trece informes las referencias aparecen solo como un listado de bibliografía al final del texto, sin mención explícita en el desarrollo de los contenidos, de manera tal que no es claro para qué y cómo dichas fuentes fueron utilizadas en la construcción del informe. La cantidad de fuentes identificadas y de veces en que son referidas tiene también diferencias significativas. Así, por ejemplo, en un informe se identifican 67 notas a pie de página[50]. En otro, un listado bibliográfico con 54 fuentes[51]. En el extremo opuesto, varios informes sólo citan o incluyen una o hasta cuatro fuentes en sus referencias[52].

En definitiva, un 62% del total de informes analizados no cuentan con referencias bibliográficas. El 38% que si incluye, las utilizan con distinta intensidad y se presentan de manera diversa, no permitiendo en algunos formatos controlar adecuadamente su uso.

Otro indicador que permite apreciar diferencias importantes en el desarrollo de los informes es su extensión. Sin poder extraer conclusiones muy definitivas a partir de estos datos, ellos sugieren que estas diferencias son expresivas de una distinta comprensión acerca de cuál es el mínimo exigible a un informe escrito. Por ejemplo, en los informes tanatológicos, el promedio de extensión de los requeridos por el MP es de 4,6 páginas versus 14 en los de la DPP. Algo similar ocurre en los peritajes clínicos en donde los promedios son de 4 y 12,3 páginas respectivamente. Estas diferencias se incrementan tratándose de los informes de laboratorio en los que la extensión de los del MP es 4,1 y los de la DPP 26,3 páginas promedio.

Como se señalaba, la heterogeneidad de los informes no es un problema per se considerando algunas diferencias que naturalmente pueden existir producto de la diversidad disciplinar o de los roles representados en el proceso. En esta dirección, la doctrina chilena ha entendido que las exigencias del artículo 315 del CPP no pueden traducirse en un único formato rígido (Cáceres, 2023, p. 87; González-Wilhelm, 2018, p. 310). Sin embargo, los casos ejemplares analizados y algunos indicadores globales recopilados en la investigación sugieren que esta heterogeneidad se daría en aspectos en donde no parece existir una justificación correcta. Ella parece más bien reflejar que los informes presentan defectos o problemas que limitan o dificultan

[50] DPP n° 9 referido a un examen médico legal.

[51] DPP n° 49 referido a un peritaje de psicología forense.

[52] DPP n° 2; DPP N° 13; DPP n° 17; DPP n° 24; DPP n° 25; DPP n° 28, DPP n° 32; DPP n° 33, DPP; DPP n° 54 y DPP n° 56.

que estos documentos se constituyan en herramientas que permitan un escrutinio estricto de esta prueba. La misma doctrina que admite grados de flexibilidad en la estructura y contenidos de los informes es clara en destacar que ella está condicionada al cumplimiento de unos mínimos obligatorios (Cáceres, 2023, p. 87; González-Wilhelm, 2018, p. 310). En muchos casos, la ausencia de esos mínimos parece agravar el problema de la heterogeneidad según se verá a continuación.

4.2.2.2. Informes incompletos

Junto con la falta de uso de fuentes o referencias bibliográficas, la investigación detectó que los informes adolecen de varios componentes de importancia. Ya se explicó la relevancia de los anexos y cómo su inclusión se está transformando en un estándar mínimo exigible a nivel comparado. En la muestra en estudio solo cuatro informes (4,4%) incluyeron un anexo[53]. Algo similar ocurre tratándose de la inclusión de explicaciones relativas a las limitaciones de los resultados obtenidos o cualificaciones sobre su alcance. Se identificaron solo tres informes (3,3%) que incluyeron tal información, al menos de manera explícita[54]. La tabla 3 resume los resultados en estos tres aspectos.

Tabla 3 Presencia de contenidos más deficitarios en informes periciales revisados (n: 90)

Ítem	% Informes MP (n: 33)	% Informes DPP (n: 57)	% Total (n: 90)
incluye referencias bibliográficas	9%	54,3%	37,7%
incluye anexo	3%	5,2%	4,4%
incluye limitaciones	9%	0%	3,3%

Fuente: elaboración propia

Adicionalmente, hay contenidos en donde la carencia de información es porcentualmente menor, pero recae sobre aspectos básicos que sorprende no se registren con cierta uniformidad en todos los reportes. Por ejemplo, tratándose de las instrucciones recibidas por el mandante, un 24,2% de los informes del MP y un 5,2% de los DPP no registran tal información. En materia de identificación de los objetivos del informe, un 15,1% de los del MP y un 12,2% de la DPP no lo contienen. Finalmente, respecto a la especificación de la información revisada al elaborar el informe, en un 3% de los solicitados por el MP y en un 17,5% de los de la DPP ella no se incluye[55].

[53] MP n° 9 (planimétrico); DPP n° 19 (neurológico); DPP n° 40 (químico farmacéutico); y, DPP n° 51 (químico farmacéutico).

[54] MP n°9 (médico criminalístico); MP n° 24 (químico); y, MP n° 25 (químico).

[55] Un problema distinto se refiere al tipo de información al cual accede el experto para realizar su informe. Como ya se ha mencionado, la evidencia comparada muestra que el acceso a información irrelevante para la tarea encargada al perito podría ser una fuente de sesgos cognitivos como el sesgo de información contextual. Urrea (2024) ha manifestado preocupación por el acceso habitual de los peritos

Como se puede observar, si bien el porcentaje de elementos ausentes disminuye sensiblemente en estas categorías, no se aprecia que exista alguna diferencia de tipo disciplinar o de rol que pudiera justificar la falta de información de esta naturaleza la que, a la vez, parece clave para poder evaluar correctamente al informe.

Un elemento más complejo para evaluar, sin un conocimiento detallado de las diversas disciplinas a las que pertenecen los peritajes analizados, se refiere a la metodología usada en su construcción. Se trata, como ya ha sido visto, de una exigencia mínima crucial. En este contexto, el estudio permitió identificar la presencia o no de una sección dedicada a la metodología utilizada. Se trata de un indicador limitado, pero que al menos arroja una primera mirada sobre la materia. Considerando esta forma de medir, un 51,5% de los informes requeridos por el MP incluyeron un capítulo específico de metodología y un 48,5% no. Tratándose de los informes de la DPP un 57,9% sí lo incluyó y un 42,1% no.

Para intentar tener una imagen un poco más precisa en esta materia debido a su importancia, se analizaron los informes que no contenían una sección de metodología para identificar si ellos incluían alguna información similar o equivalente. Se detectó que varios agregaban referencias a las actividades u operaciones realizadas o practicadas por el perito en la confección de su reporte, lo que podría ser considerado como una especificación de un contenido de tipo metodológico o, al menos, una parte de este[56]. Un 18,2% de los informes requeridos por el MP y un 15,8% de los de la DPP incluyeron información de esta naturaleza. Si consideráramos que ella es equivalente a una sección de metodología, alrededor de un 30% de los informes revisados no contuvieron un capítulo o sección dedicado al tema. Se trata de un hallazgo muy preocupante debido a la importancia de este contenido para poder controlar la calidad del trabajo pericial, pero que podría explicarse, entre otras razones, por la falta de claridad que existe sobre su exigencia en las reglas del artículo 315 del CPP como ya se tuvo oportunidad de mencionar.

La situación descrita podría agravarse si es que se hiciera un análisis más profundo de los contenidos de los informes que sí cuentan con una sección de metodología. La lectura de varios de ellos muestra un desarrollo bastante menor, en un nivel de abstracción importante o simplemente como un listado de actividades realizadas, todo lo que dificulta comprender los reales alcances del método utilizado[57]. Con todo, como ya se señaló, sin conocimiento detallado de la disciplina respectiva resulta difícil hacer un juicio más profundo sobre este punto.

---

a información no necesaria para su tarea precisamente por esta razón, constituyéndose como un riesgo importante la práctica de acceder a la carpeta completa del fiscal para su confección (pp. 276-277). En esta investigación se pudo detectar se trataba de una práctica frecuente en los casos en que se realizaron especificaciones de la información revisada.

[56] Esta pareciera ser la lógica, por ejemplo, del protocolo de autopsias simples del SML (Resolución exenta n° 905 de 2021, Servicio Médico Legal).

[57] Algunos ejemplos son DPP n° 1; DPP n° 3; DPP n° 24; DPP n° 25; DPP n° 26; y, DPP n° 27.

La tabla n° 4 resume los principales hallazgos de contenidos descritos en los párrafos precedentes.

Tabla 4 Presencia de contenidos menos deficitarios en informes periciales revisados (n: 90)

Ítem	% Informes MP (n: 33)	% Informes DPP (n: 57)	% Total (n: 90)
incluye sección de metodología	51,5%	57,9%	54,4%
incluye instrucciones de mandante	75,7%	94,7%	87,7%
incluye objetivos del informe	81,8%	87,7%	85,5%
incluye información revisada	96,9%	82,4%	87,7%

Fuente: elaboración propia

Una mención aparte requiere la falta de información curricular de los expertos. Como se ha visto, la legislación chilena la exige para acreditar idoneidad del experto, pero no necesariamente debe registrarse en el informe. Con todo, pareciera ser una buena práctica a la luz de los ejemplos de derecho extranjero y propuestas comparadas incluirla. Ella es posible de ser incorporada sin grandes costos para el experto y ayudaría mucho a mejorar la calidad de la información disponible para hacer un control más profundo de esta prueba.

Tratándose de los informes requeridos por el MP, sólo un 3% contuvo alguna información curricular del experto. Esa cifra se eleva a un 26,3% en el caso de los de la DPP. El promedio es de 18%. La regla general es que la gran mayoría de los informes (el 82%) no contienen información en esta materia. La diferencia entre ambos actores podría explicarse que, tratándose de los informes del MP de personas que trabajan instituciones públicas, la interpretación jurisprudencial prevalente del artículo 321 del CPP ni siquiera haría necesario la presentación de antecedentes curriculares ya que bastaría con la pertenencia del experto a la institución pública respectiva.

Para profundizar este hallazgo, se realizó una revisión más detallada de los casos en los que hubo inclusión de referencias curriculares en el informe. Se determinó que, en todos ellos, dichas referencias eran genéricas, es decir, una mención con mayor o menor extensión de los títulos académicos o experiencia básica, pero sin ningún tipo de detalle. Por ejemplo, en el único informe del MP en donde se contuvo esta información, ella se incluye abajo de la firma en la página final del reporte, parte en que se menciona en dos líneas la identificación de un magister, de otro en candidatura (c), el número de cédula de identidad y registro de la experta en la asociación profesional respectiva[58]. Un ejemplo con más desarrollo está constituido por un informe

[58] MP n° 2. Este tipo de información mínima parece ser la regla general ya que la mayoría de los informes que consignan datos curriculares tienen información similar. Así, tratándose de informes de la DPP: DPP N° 4; DPP n° 5; DPP n° 6; DPP n° 9; DPP n° 10; DPP n° 12; DPP n° 27; DPP n° 44; DPP n° 45; DPP n° 48; DPP n° 53; y, DPP n° 55.

requerido por la DPP. En su página inicial y a modo de pie de página se incluyen 15 líneas de información sobre el experto que muestran referencias a sus títulos y actividades de especialización profesional y a algunos aspectos de la experiencia profesional[59]. No hay reporte que contenga información sobre comparecencias previas, listado de publicaciones relevantes en los años anteriores o datos sobre calificaciones individuales o acreditaciones institucionales, entre otras.

5. A MODO DE CONCLUSIÓN

En una sociedad como la actual en donde el conocimiento científico y técnico aumenta en forma exponencial, todo indica que la prueba pericial seguirá teniendo, eventualmente incrementando, un rol de importancia en el funcionamiento de los sistemas de justicia penal. Este escenario obliga a los sistemas jurídicos a tener un tratamiento más riguroso de esta evidencia que permita aprovechar sus aportes y ventajas, pero también se haga cargo de minimizar o controlar potenciales riesgos en su uso.

Un elemento central del diseño de los sistemas procesales para favorecer un control adecuado en el uso de la prueba pericial es la exigencia de un informe escrito previo. En esta dirección, la doctrina comparada y el derecho extranjero han ido elevando las exigencias de contenidos para permitir que dicho instrumento cumpla efectivamente con tal labor. La situación en Chile en cambio es compleja. La escasa doctrina que aborda esta materia también formula recomendaciones en la línea de lo sugerido en el ámbito comparado y ha construido interpretaciones de las reglas vigentes consistentes con ellas. Con todo, normativamente se mantiene una regulación que en lo esencial tiene más de 150 años de vida y que fue diseñada en un contexto procesal y cultural muy diverso. La continuidad legislativa en este punto, a pesar del cambio de sistema procesal penal, podría explicar la poca innovación en las prácticas. Así, si bien las disposiciones del CPP podrían ser interpretadas de una manera consistente con las exigencias actuales más rigurosas sobre contenidos de los informes, la evidencia empírica compilada en esta investigación, que se suma a estudios previos, sugiere que la práctica en Chile se encuentra lejos de aquello y que porcentajes no despreciables de los informes escritos presentan debilidades relevantes de información.

Las alternativas de acción son diversas. Una primera es la de hacer un esfuerzo de reconstrucción dogmática de los actuales artículos 314 y 315 del CPP de manera de incluir en sus contenidos varios aspectos que hoy no son evidentes en su lectura y que parecen ser cruciales. Se trata de una tarea que en parte ya ha realizado la doctrina nacional[60], pero todavía sin que se trate de una tendencia muy masiva ni

[59] DPP n° 1. A pie de firma con un poco más desarrollo que el promedio: DPP n° 15.

[60] Por ejemplo, se ha sostenido que todos los elementos de contenidos que sugiere debieran confirmar el mínimo de informe pericial escrito pueden ser inferidos del artículo 315 del CPP (Cáceres, 2023, p. 99).

que exista completo acuerdo en la misma. Una segunda alternativa sería promover una reforma legal que actualice las disposiciones vigentes y regule con más claridad y precisión los contenidos y oportunidad de acceso a los informes. Se trata de un camino más dificultoso no tanto por su complejidad técnica, sino más bien por la gestión política, especialmente considerando que no parece estar instalado en Chile la noción que se está frente a un problema de cierta gravedad que requería soluciones en este nivel. Desde el punto de vista técnico en cambio, los ejemplos de derecho extranjero y los estándares y propuestas revisadas entregan una base sobre la cual no parece tan complejo construir una alternativa de cambio.

Se ha propuesto también que la vía adecuada para lograr una mayor estandarización de contenidos mínimos de los informes escritos en materia penal podría lograrse por medio de la dictación de un Auto Acordado de la Corte Suprema ya que se trata de una materia que no está bien precisada y ordenada en la legislación chilena «con miras a dar buen y correcto funcionamiento para la administración de justicia» (Cáceres, 2023, p. 103). Esta sugerencia es muy interesante, aun cuando no pareciera ser un instrumento utilizado habitualmente para regular cuestiones de esta naturaleza y, por lo mismo, podría haber resistencias a avanzar en esa línea.

Una última vía de trabajo es profundizar lo que algunas instituciones, como el SML[61], han venido haciendo los últimos años en términos del desarrollo y perfeccionamiento de protocolos que abordan los contenidos de los informes escritos, incluyendo las exigencias que hemos visto son deficitarias en la práctica nacional y poco claras en las disposiciones legales vigentes. Un límite de esta alternativa es que dichos protocolos no tienen poder vinculante para informes de expertos fuera de la institución, sin perjuicio de su potencial rol orientador para todo el sistema.

Como se puede apreciar, cualquiera de estas alternativas podría contribuir a generar mejoras importantes en la práctica chilena, aun cuando presentan limitaciones y, en algunos casos, no parece realista pensar que puedan producirse en el corto plazo. Quizás un supuesto pendiente para avanzar es tener mayor consciencia y una mejor comprensión del problema existente en esta materia. Esta investigación ha pretendido contribuir a esto por medio de una revisión crítica de la legislación chilena a la luz de propuestas y ejemplos de derecho extranjero que dan cuenta de sus debilidades y también a través de aportar evidencia empírica no disponible previamente y que pretende contribuir a la construcción de un diagnóstico más preciso en el área que aún no existe.

[61] El SML ha desarrollado, por ejemplo, estándares técnicos para la elaboración de informes periciales en salud mental (González-Wilhelm y Duce, 2024, p. 6). También se ha avanzado, como ha sido visto, en el desarrollo de protocolos en el ámbito de pericias tanatológicas, por ejemplo, las autopsias estándar (Resolución exenta n° 905 de 2021, SML).

BIBLIOGRAFÍA

Aguilera, G., y Galleguillos, H. (2025). La prueba pericial: aspectos relevantes de su admisibilidad y valoración: regulación en los procedimientos reformados en Chile. Ediciones DER.

Appazov, A. (2016). Expert evidence and international criminal justice. Springer.

Arellano, J. (2017). Desafíos de la Reforma procesal Penal en Chile: análisis retrospectivo a más de una década. Centro de Estudios de Justicia de las Américas.

Blanco, R., Decap, M., Moreno, L. y Rojas, H. (2021). Litigación estratégica en juicios orales. Tirant lo Blanch.

Cerda, R. (2010). Manual de justicia penal. Librotecnia.

Choo, A. (2009). Evidence. Oxford University Press.

Cruz, A. (2023). Curso de derecho procesal penal. Thomson Reuters.

Condemarín, P. y Lara, M. (2015). Estudio sobre instrumentalización de los peritajes psicológicos en materia de delitos sexuales en el contexto de la nueva justicia penal chilena: emergencia de nuevos escenarios (documento no publicado en poder del autor).

Defensoría Penal Pública. (2020). Memoria Anual 2019.

Defensoría Penal Pública. (2025). Memoria Anual 2024.

Departamento de Estudios Defensoría Nacional. (2016). Análisis de la gestión de peritajes de la defensa penal, Estudios n° 1.

Dirección de Estudios de la Corte Suprema (DECS). (2017). Peritajes en Chile. *https://drive.google.com/file/d/1bIcZpvY5Vhvjpq3ILykCFZuLR4oCGU_6/view*

Dror, I. (2020). Cognitive and human factors in expert decision making: six fallacies and eight sources of bias, Analytical Chemistry Vol. 92, N° 12, pp. 7998-8004.

Dror, I., *et al.* (2021). Cognitive bias in forensic pathology decisions, Journal on Forensic Sciences, pp. 1751-1757.

Duce, M. (2013). La prueba pericial. Ediciones Didot.

Duce, M. (2018). Una aproximación empírica al uso y prácticas de la prueba pericial en el proceso penal chileno a la luz de su impacto en los errores del sistema. Politica Criminal, 13(25), pp. 42-103.

Duce, M. (2022). Los errores, la aplicación al caso concreto y los sesgos cognitivos de los peritos, en C. Vázquez (Coord.), Manual de prueba pericial (pp. 143-183). Suprema Corte de Justicia de la Nación, México.

Edmond, G. (2013). Expert evidence in reports and courts. Australian Journal on Forensic Sciences, 45(3), pp. 248-262.

Edmond, G. (2015). Forensic Science Evidence and the conditions for rational (jury) evaluation. Melbourne University Law Review, 39(1), pp. 77-127.

Fierro, C. y Walker, A. (2024). El recurso de nulidad ante la Corte Suprema. Tirant lo Blanch.

Forensic Science Regulator. (2020). Forensic Science Regulator Guidance. Expert report guidance FSR-G-200 issue 4. *https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/881550/Statement_and_Report_Guidance_Appendix_Issue_4.pdf*

Fundación Amparo y Justicia (2024). Diagnóstico de la investigación de homicidios de niños, niñas y adolescentes en Chile. https://amparoyjusticia.cl/biblioteca-digital/diagnostico-de-la-investigacion-de-homicidios-de-ninos-ninas-y-adolescentes-en-chile/

Garrett, B. (2011). Convicting the innocent. Harvard University Press.

González-Wilhelm, L. (2018). Prueba pericial, litigación con peritos y medicina legal. Thomson Reuters.

González-Wilhelm, L. y Duce, M. (2024). Medicina legal en Chile: la Cenicienta sin príncipe. Medwave, (24)10, pp. 1-9, *http://doi.org/10.5867/medwave.2024.10.2978*

Keane, A. y McKeowon, P. (2020). The modern law of evidence. Oxford University Press.

Horvitz, M. y López, J. (2004). Derecho procesal penal chileno (Tomo II). Editorial Jurídica de Chile.

Martire, K., Neal, T., Gobet, F., Chin, J., Berengut, J. y Edmond, G. (2025). Psychological insights for judging expertise and implications for adversarial legal contexts. Nature Reviews Psychology, 4, pp. 264-276.
Maturana, C. y Montero, R. (2010). Derecho procesal penal (Tomo II). Abeledo Perrot.
Meza, M. (2024). El informe pericial químico en la determinación de residuales de disparos. Contenidos mínimos, razonamiento experto y debido proceso. Revista Quastio Facti, (6), pp. 285-298.
Morgan, J. (2023a). Wrongful convictions and forensic science errors. CRC Press.
Morgan, J. (2023b). Wrongful convictions and claims of falso or misleading forensic evidence. Journal of Forensic Sciences, 00, pp. 1-54.
National Commission on Forensic Science (NCFS). (2015). Views of the Commission: documentation, case record and reports contents. *https://www.justice.gov/ncfs/file/818191/dl?inline*
National Commission on Forensic Science (NCFS). (2016). Recommendations to the Attorney General pretrial Discovery. *https://www.justice.gov/ncfs/file/865011/dl?inline*
National Research Council of the National Academies of Science (NRC-NAS). (2009). Stregngthening forensic sciences in United States: a path forward. The National Academies Press.
Neal, T., Liener, P., Denne, E. y Singh, J. (2022). A general model of cognitive bias in human judgment and systematic review specific to forensic mental health. Law and Human Behavior, 46(2), pp. 99-120.
Paccioco, D. (2009). Unplugging jukebox testimony in an adversarial system: strategies for changing the tune on partial experts. Queen's Law Journal, 34(2), pp. 565-610.
Quagley-Macbride, A., DROR, I., Roy, T., Garret, B. y Kukuca, J. (2022). A practical tool for information management in forensic decisions: using linear sequential unmasking-expanded (LSU-E) in casework. Forensic Science International: Sinergy (Vol. 4), pp. 1-6.
Quezada, J. (1994). Tratado de derecho procesal penal. Editorial Jurídica ConoSur.
Quijada, D. (2011). Aportes al debido proceso penal en la evaluación pericial psicológica y psiquiátrica a imputados de delitos sexuales: ¿existe relación de coherencia entre solicitud y respuesta pericial?, en C. Fuentes (Coord.), Diez años de la reforma procesal penal en Chile (pp. 457-499). Ediciones Universidad Diego Portales.
Resolución exenta n° 905 de 2021 [Servicio Médico Legal]. 20 de abril de 2021
Resolución exenta n° 1501 de 2021[Servicio Médico Legal]. 19 de julio de 2021
Summersby, S., Edmond, G., Kemp, R., Ballantyne, K. y Martire, K. (2024). The effect of following best practice reporting recommendations on legal and community evaluations of forensic examiners reports. Forensic Science International, (359), pp. 1-14.
The Law Commission. (2011) Expert evidence in criminal proceedings in England and Wales. London Stationery Office.
Urrea, J. (2024). Contenido, forma y valoración racional de la prueba pericial. Revista Quastio Facti, (6), pp. 271-284.
Vázquez, C. (2022). Guía sobre el contenido de los informes periciales y su impacto en el debido proceso. Consejo Superior de la Judicatura.
Vázquez, C. (2024). A manera de introducción a la sección a propósito del contenido de los informes periciales. Revista Quastio Facti, (6), pp. 251-256.
Vredeveldt, A., Van Rosmalen, E., Van Koppen, P., Dror, I. y Otgaar, H. (2024). Legal psychologists as experts: Guidelines for minimizing bias. Psychology, Crime, & Law (Vol. 30, N° 7), pp. 705-729.

Jurisprudencia Corte Suprema:

Corte Suprema. Rol n° 5597-2009;2 de noviembre de 2009.
Corte Suprema. Rol n° 2866-2012; 17 de junio 2013.
Corte Suprema. Rol n° 4554-2014; 10 de abril de 2014.
Corte Suprema. Rol n° 11515-2014; 21 de julio de 2014.
Corte Suprema. Rol n° 448-2015; 11 de marzo de 2015.

Corte Suprema. Rol n° 4712-2017; 28 de marzo de 2017.
Corte Suprema. Rol n° 12646-2019; 25 de junio de 2019.
Corte Suprema. Rol n° 97049-2021; 29 de diciembre de 2021.
Corte Suprema. Rol n° 246257-2023; 9 de enero de 2024.
Corte Suprema. Rol n° 14553-2024; 15 de julio de 2024.
Corte Suprema. Rol n° 16613-2024; 19 de julio de 2024.
Corte Suprema. Rol n° 32047-2024; 18 de octubre de 2024.
Jurisprudencia Cortes de Apelaciones:
Corte de Apelaciones de Iquique. Rol n° 78-2004; 26 de noviembre de 2004.
Corte de Apelaciones de Valdivia. Rol n° 263-2005; 2 de enero de 2005.
Corte de Apelaciones de Chillán. Rol n° 93-2007; 8 de agosto de 2007.
Corte de Apelaciones de Antofagasta. Rol n° 276-2010; 3 de septiembre de 2009.
Corte de Apelaciones de Santiago. Rol n° 2130-2009; 21 de octubre de 2009.
Corte de Apelaciones de Arica. Rol n° 145-2010; 7 de julio de 2010.
Corte de Apelaciones de Valparaíso. Rol n° 865-2010; 8 de octubre de 2010.
Corte de Apelaciones de Coyhaique. Rol n° 42-2023; 21 de abril de 2023.
Corte de Apelaciones de Antofagasta. Rol n° 877-2023; 3 de julio de 2023.
Corte de Apelaciones de Talca. Rol n° 830-2023; 18 de agosto de 2023.

Quaestio facti. Revista Internacional sobre Razonamiento Probatorio
Quaestio facti. International Journal on Evidential Legal Reasoning
Sección: Ensayos
2025 I 10 pp. 283-325
Madrid, 2026
DOI: 10.33115/udg_bib/qf.i10.23200
Marcial Pons Ediciones Jurídicas y Sociales
© Sergio Cruz Arenhart
ISSN: 2604-6202
Recibido: 13/05/2025 | Aceptado: 03/11/2025 | Publicado *online*: 12/01/2026
Editado bajo licencia Reconocimiento 4.0 Internacional de Creative Commons

LA CONDUCTA DE LAS PARTES COMO PRUEBA EN EL PROCESO CIVIL BRASILEÑO

Sergio Cruz Arenhart

RESUMEN: El texto explora la conducta de las partes como elemento probatorio en el proceso civil brasileño. Se examina la tendencia de la legislación brasileña a utilizar esta prueba solamente como indicio, lo que plantea varios obstáculos para su empleo en el ámbito del proceso civil. Se critica este punto de vista con el fin de ampliar el uso de la conducta de las partes como medio de prueba, ya sea en combinación con otras pruebas o de forma aislada (sobre todo en casos como la tutela preventiva y la protección de derechos fundamentales). Por último, se sugieren algunos parámetros para su uso y control, como la motivación adecuada y la preservación del contradictorio.

PALABRAS CLAVE: conducta de las partes; prueba indiciaria; suficiencia probatoria; tutelas inhibitorias; bienes jurídicos involucrados; justificación adecuada y racional de la decisión; contradictorio.

THE BEHAVIOR OF PARTIES AS EVIDENCE IN BRAZILIAN CIVIL PROCEDURE

ABSTRACT: The text explores the conduct of the parties as an element of evidence in Brazilian civil proceedings. It examines the tendency of Brazilian law to use this evidence only as an circumstantial evidence, which poses several obstacles to its use in civil proceedings. This point of view is criticised in order to expand the use of the parties' conduct as a means of evidence, either in combination with other evidence or in isolation (especially in cases such as preventive protection and the protection of fundamental rights). Finally, some parameters for its use and control are suggested, such as adequate motivation and the preservation of the adversarial dialogue.

KEYWORDS: the behaviour of parties; circumstantial evidence; evidentiary sufficiency; inhibitory injunctions; legal interests involved; adequate and rational justification of decisions; adversarial proceedings.

SUMARIO: 1. ENTRE BÚSQUEDAS PERSONALES Y TESTIMONIOS FIABLES.— 2. EN BÚSQUEDA DE UNA TIPOLOGÍA DE LOS COMPORTAMIENTOS RELEVANTES PARA EL DERECHO PROBATORIO.— 3. EFECTOS *LEGALMENTE* PREVISTOS PARA CIERTAS CONDUCTAS PROCESALES Y CLÁUSULAS GENERALES DE INTERPRETACIÓN DE ESAS CONDUCTAS.— 4. EL DERECHO BRASILEÑO Y LA INTERPRETACIÓN GENERAL DE LA CONDUCTA DE LAS PARTES PARA FINES PROBATORIOS.— 5. LA CALIDAD DE LA PRUEBA DERIVADA DEL COMPORTAMIENTO DE LAS PARTES.— 6. MÓDULOS DE PRUEBA, CATEGORÍAS DE HECHOS Y LA PRUEBA POR LA CONDUCTA DE LOS SUJETOS.— 7. ¿HAY ALGUNA CONCLUSIÓN POSIBLE?.— 8. REFERENCIAS BIBLIOGRÁFICAS

1. ENTRE BÚSQUEDAS PERSONALES Y TESTIMONIOS FIABLES

El artículo 244 del Código de Procedimiento Penal brasileño establece que la búsqueda personal está permitida, independientemente de una orden judicial, entre otros casos, siempre que haya una "sospecha fundada" de que la persona tenga en su posesión un arma prohibida u otros objetos que puedan constituir el cuerpo del delito. La noción de lo que constituye esta "sospecha fundada" —que autoriza el uso de la medida— es objeto de interminable debate en la doctrina y jurisprudencia brasileñas, así como en el ámbito internacional. Por regla general, para satisfacer este requisito se exige que existan "elementos objetivos" que indiquen que la persona presenta una actitud sospechosa, no siendo suficiente la mera impresión subjetiva del agente público para autorizar la búsqueda personal. Aunque hasta este punto la cuestión es razonablemente tranquila, la noción de lo que son estos "elementos objetivos" puede variar inmensamente. Así, por ejemplo, la Corte Interamericana de Derechos Humanos, al juzgar el caso *Fernández Prieto y Tumbeiro vs. Argentina* (decisión del 1 de septiembre de 2020), entendió que la mera circunstancia de que la persona pareciera nerviosa, estuviera vestida de manera extraña o presentara respuestas inconsistentes no calificaba estos elementos objetivos como para justificar la búsqueda personal.[1]

[1] «78. La Corte recuerda que, en su declaración de 15 de enero de 1998, uno de los agentes intervinientes en la detención del señor Tumbeiro narró las circunstancias que llevaron a la misma expresando que ese día, cuando se encontraba recorriendo el "radio jurisdiccional", pudo "observar a una persona del sexo masculino el cual se hallaba vestido con zapatos negros, pantalón jean azul y camisa a cuadros, el que al observar la presencia policial se mostró sumamente nervioso y dubitativo a la vez que intentaba eludir el paso del móvil policial. Atento a ello, se procedió a detener su marcha y con el objeto de verificar si registraba algún impedimento legal, se invitó al mismo a ascender al móvil hasta tanto se determine mediante sistema digito radial su identidad". El agente expresó que, "debido a que continuaba sumamente nervioso, se solicitó la cooperación de los testigos […] junto a los cuales se procedió a examinar los efectos personales de dicha persona". Esta declaración, concatenada a la relación fáctica que figura en las sentencias del

En los tribunales brasileños, también se consideró que el mero nerviosismo[2] o el simple hecho de que alguien vistiera ropa que pudiera ocultar armas es insuficiente para autorizar tal medida.[3] Por otro lado, y en un sentido claramente diferente, ya se ha dicho que la circunstancia de que alguien demuestre aprensión, mirando insistentemente hacia atrás, sería suficiente para satisfacer dichos criterios objetivos.[4] También se ha considerado que el hecho de que alguien esté en un lugar sospechoso y, al constatar la presencia policial, altere radicalmente su comportamiento, desechando un envoltorio, es suficiente para satisfacer el criterio legal.[5]

Decisiones tan diversas sobre la misma cuestión demuestran la dificultad que se tiene al interpretar lo que se puede extraer del comportamiento humano. Muchas veces, debido a sesgos cognitivos, preconceptos o criterios subjetivos, alguien puede considerar sospechosa cierta conducta, mientras que para otra persona ese mismo comportamiento puede sonar absolutamente normal. Al emitir un voto disidente en el *Habeas Corpus* n. 169.788/SP,[6] que trataba sobre las búsquedas domiciliarias, el

---

proceso penal, permite a la Corte advertir que el señor Tumbeiro fue detenido para que se identificara en virtud de tres hechos: a) se mostró nervioso ante la presencia de los policías; b) no estaba vestido conforme al modo de vestir percibido por los agentes como propio de la zona por la que transitaba, y c) contestó que se encontraba buscando un material "totalmente extraño a lo que podía obtenerse en los comercios aledaños". 79. La Corte nota que, de conformidad con la Ley 23.950, la retención temporal con fines de identificación debe estar debidamente fundada en circunstancias que "hagan presumir que alguien hubiese cometido o pudiere cometer algún hecho delictivo o contravencional". En ese sentido, en el caso concreto, el Tribunal considera que ninguna de las razones que dio la policía para retener al señor Tumbeiro y solicitarle su identificación constituían en sí mismas, o en conjunto, hechos o informaciones suficientes y concretas que permitan a un observador razonable inferir objetivamente que probablemente había cometido o estaba por cometer un hecho delictivo o contravencional. Por el contrario, las razones que motivaron la detención con fines de identificación del señor Tumbeiro parecieron responder a preconceptos sobre como debe verse una persona que transita en un determinado lugar, cómo debe comportarse ante la presencia policial, y qué actividades debe realizar en ese lugar. (...). 81. Ante la ausencia de elementos objetivos, la clasificación de determinada conducta o apariencia como sospechosa, o de cierta reacción o expresión corporal como nerviosa, obedece a las convicciones personales de los agentes intervinientes y a las prácticas de los propios cuerpos de seguridad, lo cual comporta un grado de arbitrariedad que es incompatible con el artículo 7.3 de la Convención Americana. Cuando adicionalmente estas convicciones o apreciaciones personales se formulan sobre prejuicios respecto a las características o conductas supuestamente propias de determinada categoría o grupo de personas o a su estatus socio-económico, pueden derivar en una violación a los artículos 1.1 y 24 de la Convención» (decisión disponible en https://www.corteidh.or.cr/docs/casos/articulos/seriec_411_esp.pdf, accedido en juño de 2025).

[2] STJ, 6ª Turma. AgRg no REsp 2.181.502/SC. Rel. Min. Rogério Schietti Cruz. DJe 15.04.25; STJ, 6ª Turma. HC 659.689/DF. Rel. Min. Sebastião Reis Junior, DJe 18.06.21.

[3] STF, 1ª Turma. HC 81.305/GO. Rel. Min. Ilmar Galvão. DJe 22.02.02.

[4] «Segundo se infere, a atuação policial se originou de denúncia de um terceiro (não identificado) sobre a prática de traficância no local, por duas pessoas em uma motocicleta, e o ora agravante, que estava na garupa, demonstrou apreensão, ao perceber a aproximação da guarnição, olhando constantemente para trás, e trazia algo na cintura, posteriormente onde foi localizada a droga e os objetos para embalo/venda dos entorpecentes. Contexto suficiente a indicar atitude concretamente suspeita do réu para permitir a abordagem» (STJ, 5ª Turma. AgRg no HC 958.143/MG. Rel. Min. Ribeiro Dantas. DJe 08.04.25).

[5] STJ, 5ª Turma. AgRg no AgRg no REsp 2.141.306/PR. Rel. Min. Joel Ilan Paciornik. DJe 25.04.25.

[6] STF, Pleno. HC 169.788. Redactor del fallo Min. Alexandre de Moraes. Dje 04.03.24.

Ministro Edson Fachin llegó a referirse, como expresión de ese móvil subjetivo, a un "sexto sentido" de la autoridad policial, inservible para sustentar cualquier otra medida instructoria. La subjetividad en la evaluación del "comportamiento sospechoso", así, puede llevarnos a reflexionar si la actitud de alguien en ciertas circunstancias puede o no ser tomada como elemento relevante para fines probatorios en el proceso.

Si la valoración del comportamiento individual ya reúne puntos sensibles que ponen en duda la viabilidad de su empleo en el proceso, imagínese adoptar comportamientos pasados de alguien —a veces, comportamientos aislados; a veces, repetidos— como argumento de prueba en el proceso.

En septiembre de 2021, el Superior Tribunal de Justicia brasileño afectó para juzgamiento como materia repetitiva la siguiente cuestión: "definir la posibilidad de imposición de tutela inhibitoria, así como de responsabilidad civil por daños materiales y morales colectivos causados por el tráfico con exceso de peso en carreteras".[7] El tema adquirió importancia porque, en diversos estados del país, el Ministerio Público Federal había presentado acciones colectivas contra empresas de transporte, solicitando la prohibición de que permitieran la salida de vehículos de carga con exceso de peso y la condena de esas empresas al pago de daños materiales y morales colectivos. Se alegaba, en líneas generales, que esas empresas habían sido reiteradamente sancionadas, en la esfera administrativa, por permitir que sus vehículos de carga transitaran por las carreteras nacionales con exceso de carga. Se decía que este comportamiento perjudicaba sobremanera la vida útil de las carreteras, además de poner en riesgo a personas y otros valores sociales relevantes. Por ello, y considerando que la sanción administrativa se mostraba insuficiente para cohibir la práctica ilegal, se solicitaba judicialmente la adopción de otras medidas más severas de carácter inhibitorio, así como la indemnización de los daños causados.

En la gran mayoría de los casos, los tribunales locales rechazaron estas acciones, afirmando que las eventuales infracciones sobre el límite de carga tolerado ya eran sancionadas administrativamente, mediante multas de tráfico, no siendo razonable que las empresas fueran nuevamente castigadas por vía judicial.[8] En varios de los casos, se añadió a este fundamento también la inexistencia de prueba del nexo de causalidad (entre la infracción cometida por las empresas y los daños causados en las carreteras), ya que muchos otros factores podrían contribuir a estos daños, no pudiéndose atribuir a la conducta reiterada de las empresas un protagonismo o papel determinante que pudiera atraerles la responsabilidad por los perjuicios.[9]

---

[7] Tema 1.104 (STJ, 1ª Seção. ProAfR no REsp 1.908.497/RN. Rel. Min. Assusete Magalhães. DJe 10.09.21 e STJ, 1ª Seção. ProAfR no REsp 1.913.392/MG. Rel. Min. Assusete Magalhães. DJe 10.09.21).

[8] V.g., TRF1, 6ª Turma. AI 0056520-92.2012.401.0000/DF. Rel. Desembargador Federal Carlos Moreira Alves. DJe 23.08.13; TRF4, 3ª Turma, AC 5001629-36.2015.404.7214. Rel. Desembargador Federal Rogério Favreto. DJe 08.02.18.

[9] Con ese argumento, por ejemplo, subrayó la Jueza Federal Marga Tessler, al examinar un recurso de apelación sobre la cuestión que «qualquer estatística que se analise vai indicar como principais causas dos acidentes nas rodovias brasileiras a imprudência e o excesso de velocidade, de modo que os demais

Ante el volumen de causas con discusiones similares (y también porque existía un entendimiento de aquella Corte Superior sobre el tema en sentido contrario), el Superior Tribunal de Justicia consideró relevante examinar la cuestión, a fin de generar un entendimiento vinculante para todo el país. Y, al juzgar la materia, el Superior Tribunal de Justicia concluyó por fijar la tesis vinculante de que "el derecho al tránsito seguro, así como los notorios e inequívocos daños materiales y morales colectivos derivados del tráfico reiterado, en carreteras, de vehículos con exceso de peso, autorizan la imposición de tutela inhibitoria y la responsabilización civil del agente infractor".[10] Entre los varios argumentos empleados para llegar a esta conclusión, la Corte subrayó que es *notorio* el impacto del transporte con exceso de peso en el deterioro de las carreteras nacionales. Se consideró que numerosos análisis puntuales reafirman esta contribución,[11] lo que permite concluir que este transporte ilegal

---

cidadãos usuários da rodovia também contribuem para a violação dos itens que são a causa do pedido indenizatório. Mas não é só. A geometria das rodovias brasileiras, a ausência de sinalização adequada e barreiras de proteção, curvas com forças centrífugas mal projetadas também o são e estas são responsabilidade do Estado. Além dos fatores acima referidos, outros dois pontos, de interesse mais direto na presente demanda, merecem ser citados, quais sejam: a ausência de fiscalização adequada e a qualidade da pavimentação asfáltica. A fiscalização ineficiente, inclusive para fins de cumprimento da vedação de circulação com veículos com excesso de peso, é um fator importante para descumprimento da regra. À vista destes elementos e das demais questões abordadas na presente decisão, conclui-se que a conduta da ré, individualmente considerada, não é suficiente para causar dano à coletividade, porquanto ela está inserida num contexto em que diversas situações convergem para o problema, sem que se possa destacar uma como de maior relevância.» (TRF4, 3ª Turma. AC 5018457-33.2016.404.7001. Rel. Desembargadora Federal Marga Barth Tessler. DJe 03.06.20).

[10] STJ, 1ª Seção. REsp 1.908.497/RN. Rel. Min. Teodoro Silva Santos. DJe 04.12.24; STJ, 1ª Seção. Resp 1.913.392/MG. Rel. Min. Teodoro Silva Santos. DJe 04.12.24. El fallo aún no es definitivo, aunque ya se hayan rechazado otros recursos opuestos por las empresas demandadas (es posible acompañar la tramitación del proceso en el siguiente link: https://processo.stj.jus.br/processo/pesquisa/?tipoPesquisa=tipoPesquisaNumeroRegistro&termo=202003215737).

[11] «Estudos revelam que a prática dessa infração administrativa é corriqueira nas rodovias do País. Segundo o "Atlas da Década de Ações para Segurança Viária", que traçou um panorama geral das condições das rodovias federais entre os anos de 2011 a 2020, o excesso de carga está entre as causas presumíveis dos acidentes de trânsito ocorridos no País, motivo pelo qual foram desenvolvidas diversas ações para coibir a prática, a exemplo "Operação de Fiscalização de Peso e Dimensões - Operação Carga Pesada". Apenas a título ilustrativo, no ano de 2020 foi noticiado que a Polícia Rodoviária Federal registrou um aumento de 18,4% de flagrantes por excesso de peso em veículos de carga nas rodovias federais do Piauí» (https://www.gov.br/prf/pt-br/noticias_anteriores/estaduais/piaui/2021/Marco%202021/carga-pesada-prf-registra-em-2020-um-aumento-de-18-4-nos-flagrantes-de-excesso-de-peso-em-veiculos-de-carga-nas-rodovias-federais-do-piaui). Mais recentemente, em março de 2022, uma autuação em Vitória da Conquista, na Bahia, resultou na apreensão de um veículo de carga com mais de 40 toneladas de excesso (https://www.gov.br/prf/pt- br/noticias_anteriores/estaduais/bahia/marco-2022/com-mais-de-40-toneladas-de-excesso-de-peso-bi-trem-e-apreendido-pela-prf-na-br-116-em-vitoria-da conquista). Os dados de sinistralidades de trânsito registrados pela Polícia Rodoviária Federal entre os anos de 2019 a 2024 nas rodovias federais revelam que a principal causa dos acidentes relaciona-se a defeitos na via. Foram identificados, ao longo do quinquênio, 8.195 defeitos nas rodovias, que lesionaram 10.228 pessoas e levaram a óbito outras 399. (...) Por sua vez, relatório de auditoria elaborado em 25/05/2012 pelo Tribunal de Contas da União (Processo nº 014.206/2012-9) apontou que o excesso de peso acelera os processos de degradação do pavimento de modo exponencial. O estudo tomou como base o experi-

contribuye significativamente a los daños en disputa.[12] Serían así notorios los daños derivados del comportamiento irregular de las empresas, lo que dispensaría prueba y acarrearía la responsabilidad civil reclamada.[13]

Para el punto que aquí interesa, debe subrayarse que lo que hizo la Corte fue considerar que el comportamiento reiterado de las empresas de transporte generaba una especie de *notoriedad* en el perjuicio causado a las carreteras brasileñas y, en consecuencia, autorizaría su responsabilidad civil —además de la adopción de medidas que evitaran la repetición del comportamiento ilegal. Sin excluir que otros factores puedan también colaborar en ese deterioro de las carreteras, el Tribunal concluyó que este comportamiento insistente justificaba la adopción de medidas judiciales contra esas empresas.[14]

---

mento realizado pela American Association of State Highway Officials - AASHO, o qual concluiu que o incremento de 10% no sobrepeso reduz a vida útil do pavimento em 40% (Lei da Quarta Potência). Portanto, não há dúvidas que a operação de transporte com sobrecarga é extremante deletéria, uma vez que provoca a deterioração acelerada da pavimentação, interfere no regular desempenho dos veículos, compromete a segurança viária, afeta o meio ambiente e desequilibra o regime de concorrência."

[12] «Do explanado, não pairam dúvidas de que os danos causados pela conduta do agente que reiteradamente trafega com excesso de carga nas rodovias acarreta, em claro nexo de causalidade, prejuízos de ordem material ao patrimônio público e moral a toda coletividade. Conforme amplamente demonstrado, o transporte de cargas com excesso de peso maximiza a produtividade e o lucro do transportador em detrimento do interesse público na preservação da pista de rolamento, na manutenção da segurança viária, na preservação do meio ambiente equilibrado e na escorreita observância do regime de concorrência.»

[13] «Conforme visto, a conduta consciente de transitar com veículo com sobrepeso nas rodovias gera prejuízos materiais a bem de uso comum do povo, haja vista a deterioração e redução da vida útil da pavimentação, que gera elevado dispêndio de recursos necessários à recuperação da via pública. No que se refere à necessidade de demonstração efetiva dos danos materiais, por tudo o que já foi dito, conclui-se pela aplicabilidade do disposto no art. 374, I, do CPC/2015, segundo o qual não dependem de prova os fatos notórios. Por fato notório compreende-se aquele de caráter evidente e conteúdo universal, visto que "encerram verdades históricas, científicas ou geográficas de reconhecimento geral. Desta sorte, o que é notório não reclama prova —notoria non egent probationem—, porque a ninguém é lícito desconhecê-lo. Essa notoriedade geral implica que o fato seja do conhecimento de toda a coletividade independentemente de sua publicidade" (FUX, Luiz. Curso de Direito Processual Civil. Grupo GEN, 2022).»

[14] Solo para contrastar con esta conclusión, cabe aún reflexionar sobre la construcción de las razones de esta responsabilidad con los motivos que llevaron al Min. Raul Araújo, en decisión unipersonal [*monocrática*], a rechazar la responsabilidad civil de una empresa farmacéutica en otro caso. La cuestión trataba de una acción de reparación de daños, interpuesta por un individuo que había utilizado el medicamento Celebra y, posteriormente, fue diagnosticado con infarto de miocardio. Alegaba, con base en un estudio realizado en Estados Unidos (que indicaba el aumento del riesgo de infarto por el uso de este medicamento), que existía nexo de causalidad, lo que permitiría la responsabilidad de la empresa farmacéutica. Rechazada la pretensión en 1º y 2º grados de jurisdicción, también en el Superior Tribunal de Justicia la conclusión fue la misma. Señaló el Min. Raul Araújo, al examinar el recurso, que "la prueba estadística, pura y simplemente, es inservible para la comprobación del nexo causal, por cuanto fundada en razonamiento inductivo sujeto a innumerables variaciones que no se compatibilizan con la individualización necesaria de la causalidad. Solo la hipótesis de frecuencia muy próxima a la efectiva ocurrencia (100%) es que tornaría relevante la prueba estadística en el análisis causal". Aunque el argumento, aquí, se centró en la cuestión de la prueba estadística, es interesante observar que, aquí, el argumento de las concausas fue considerado como relevante para excluir la responsabilidad de la empresa, sobre todo porque no había indicación precisa sobre la conexión entre el uso del medicamento y el perjuicio (infarto) ocurrido. V. STJ, AREsp 1.581.491/PR. Rel. Min. Raul Araújo. DJe 29.08.23

Todo esto pone en evidencia la importancia que el comportamiento —pasado o actual— de alguien debe tener en el ámbito probatorio. Sin duda, el comportamiento, los hábitos y el historial pasado de los sujetos se emplean usualmente en diversos campos. Sirve, por ejemplo, para que podamos elegir lugares, alojamientos y compras (mediante las calificaciones y comentarios existentes en plataformas como AirBNB, Amazon o TripAdvisor). Sirve para que las empresas de seguros bonifiquen o aumenten los valores cobrados por la renovación de seguros de vehículos o residenciales. Sirve para que las empresas de marketing puedan evaluar estrategias que optimicen la publicidad de bienes y servicios (*targeted advertising*). Sirve, en fin, para innumerables decisiones del mercado, que necesita trabajar con tendencias y hábitos de comportamiento.[15]

También no hay duda de que el Derecho, habitualmente, atribuye importancia a la conducta pasada de las personas, para diversos fines. Recuérdese que el Derecho Penal confiere un valor significativo a la conducta pasada del agente. La conducta social y los antecedentes del agente y la conducta de la víctima son elementos expresamente tomados en consideración por la ley para la fijación de la pena en caso de comisión de delitos (art. 59, del Código Penal brasileño). Además, la reincidencia, es decir, la comisión de otro delito después de su condena por un delito anterior, es un elemento que, también por imposición legal, debe siempre agravar la pena penal (art. 61, inc. I, del Código Penal brasileño). También en el plano no penal, las conductas anteriores son un elemento relevante para el derecho, sirviendo, por ejemplo, para la interpretación de la voluntad en los negocios jurídicos,[16] como elemento para la adquisición de propiedad por usucapión,[17] como causa para el divorcio[18] o para

---

[15] Como observa Schauer, «often in such cases, a common argument against using someone's previous behavior as evidence for what they might have done now is that what is past is past. Just because someone has done something before does not mean that they will necessarily do it again, and especially under different circumstances, so the argument goes. But would those who make that argument accept a late-night ride from someone who had been convicted on three previous occasions of driving while intoxicated, or, criminal convictions aside, from someone they knew to have been involved in multiple automobile accidents? (…) Past acts are commonly considered to be relevant in predicting *future* behavior, which is why convicted child molesters do not get hired as camp counselors or TSA agents, and why confessed plagiarists rarely secure positions as college professors» (Schauer, 2022, p. 207-208).

[16] Dice, por ejemplo, el artículo 113, del Código Civil brasileño, que «Os negócios jurídicos devem ser interpretados conforme a boa-fé e os usos do lugar de sua celebração». Y el párrafo 1º subraya que «§ 1º A interpretação do negócio jurídico deve lhe atribuir o sentido que: I - for confirmado pelo comportamento das partes posterior à celebração do negócio; II - corresponder aos usos, costumes e práticas do mercado relativas ao tipo de negócio (...)».

[17] De hecho, el *animus domini*, o sea, la posesión de bien «como suyo», es elemento relevante para la fijación de los plazos para la adquisición de bienes por usucapión (v.g., artículos 1.238 a 1.240, del CC brasileño).

[18] «Artículo 1.573. Podem caracterizar a impossibilidade da comunhão de vida a ocorrência de algum dos seguintes motivos: I - adultério; II - tentativa de morte; III - sevícia ou injúria grave; IV - abandono voluntário do lar conjugal, durante um ano contínuo; V - condenação por crime infamante; VI - conduta desonrosa.»

la imposición de un acuerdo de colaboración [*leniência*] a empresas que violaron las reglas de protección de la orden económica. [19]

Desde otro ángulo, es cierto que, para el derecho procesal, la conducta de las partes es muy relevante, exigiéndoles la buena fe (art. 5º, del CPC brasileño) y la colaboración (art. 6º, del CPC brasileño) en la búsqueda de una solución justa, efectiva y oportuna de la controversia. Se les imponen innumerables deberes (art. 77, del CPC brasileño) y se combate su comportamiento ilícito (arts. 80 y 774, del CPC brasileño). Hay, por lo tanto, una clara preocupación de la ley con la conducta de los sujetos parciales del proceso y, en especial, del proceso civil.

No obstante, ¿puede la conducta de estos sujetos tener una amplia relevancia probatoria en el ámbito del derecho procesal? ¿Es posible que el juez concluya la procedencia o improcedencia de cierta pretensión *basándose únicamente en el comportamiento pasado de los sujetos parciales del proceso o, quizás, en un único comportamiento pasado de alguno de ellos*?

Desde ya, se advierte que no es finalidad de este estudio examinar el deber de veracidad atribuido a las partes, tal como se mencionó, en la aportación de pruebas o en la declaración sobre los hechos del proceso. El objetivo aquí es examinar si el comportamiento «general» de las partes, su conducta —externa al proceso e incluso interna, en su relación con los actos del proceso— puede ser interpretada como prueba para la formación de la convicción judicial. [20]

2. EN BÚSQUEDA DE UNA TIPOLOGÍA DE LOS COMPORTAMIENTOS RELEVANTES PARA EL DERECHO PROBATORIO

Para empezar, parece importante considerar de modo general los posibles comportamientos que repercuten en el proceso, con el objetivo de delimitar mejor el objetivo de este texto. Es cierto que, en sentido amplio, *todo comportamiento* puede ser relevante para el proceso y estudiarlos todos sería una empresa demasiado vasta para los límites de un pequeño estudio. Así, es importante agrupar estas conductas en grandes grupos, con el objetivo de aclarar el objetivo específico de este análisis.

---

[19] V.g., artículo. 86, de la Ley 12.529/11.

[20] Como advierten Carmen Vázquez y Mercedes Fernándes López, «si hay algo que caracteriza la prueba en el ámbito jurídico es que *no* necesariamente tomamos las decisiones fácticas con toda la información relevante disponible. Ello es explicable porque en el derecho no solo tenemos el objetivo de determinar correctamente los hechos, sino de hacerlo de manera tal que sea compatible con otros derechos de las personas. La discusión, entonces, debería centrarse en a cuáles de las reglas de exclusión que actualmente tienen nuestros sistemas jurídicos, o el mexicano en particular, tienen sentido en el balance entre los objetivos de aplicar las consecuencias jurídicas solo cuando realmente hayan pasado los hechos previstos en esas normas y la manera en que podemos respetar otros derechos» (Vázquez i López, 2022, p. 160)

a) Las conductas judiciales y las conductas extrajudiciales. Por un lado, parece evidente que se pueden clasificar los comportamientos de las partes según ocurran dentro o fuera del proceso. Si las conductas externas a la relación procesal pueden ser relevantes para el proceso, las conductas ocurridas en el curso del proceso tienen, naturalmente, consecuencias para el análisis de aquello que es sometido a la apreciación judicial.

La legislación procesal, por cierto, comúnmente regula comportamientos similares, dándoles un régimen distinto según ocurran dentro o fuera del proceso. Es lo que se observa, por ejemplo, en el tratamiento dispensado a la confesión *judicial* (arts. 390 a 393 del CPC brasileño) y a la confesión *extrajudicial* (art. 394 del CPC brasileño) o en la disciplina ofrecida a la autocomposición *judicial* (art. 515, II del CPC brasileño) y *extrajudicial* (arts. 515, III y 784, inc. IV, del CPC brasileño).

No obstante que la legislación atribuya, muchas veces, diversas consecuencias a estos comportamientos en razón del *locus* en que estos actos ocurren (en juicio o fuera de él), no parece que, para fines probatorios, haya gran relevancia en esta distinción. Si bien es cierto que la circunstancia de que esta conducta haya ocurrido en juicio facilita su constatación y, así, se alinea con una de las principales finalidades de la *inmediatez* en el proceso, se trata ahí simplemente de una cuestión de facilidad de acceso a la prueba de la conducta. El valor probatorio de la conducta, sin embargo, puede ser el mismo si ocurre dentro o fuera del proceso.

Por otra parte, considerando que la existencia de un régimen específico, atribuido por ley, a algunos de estos comportamientos —muchas veces con la intención de retirar de la esfera de la prueba, por medio de ficciones jurídicas, ciertas afirmaciones de hecho, que acaban siendo consideradas por el legislador fuera del ámbito de la prueba (v.g., art. 374 del CPC brasileño)—, su estudio implicaría un número infinito de cuestiones puntuales, que desbordan los límites de un análisis inicial, tal como aquí se pretende. Por todo ello, este criterio no será relevante para el presente estudio.

b) Las conductas personales y las conductas practicadas por medio de un tercero (abogado). También se podría pensar en observar las conductas de las partes desde la óptica de quien las realiza efectivamente. Considerando que, en el proceso civil, para la práctica de actos procesales se exige, por regla general, la asistencia de un profesional del Derecho (asimilable, en términos generales, a la idea del abogado), sería posible distinguir estos comportamientos según sean practicados *por medio de abogados* o por la propia parte de manera directa.

Nuevamente, el derecho procesal da relevancia a esta diferencia en muchas circunstancias, aludiendo a los «actos procesales personalísimos». Así como ocurre con los actos (materiales) personalísimos —a ejemplo de la celebración de un testamento público o la renuncia a derechos personalísimos—, estos son los actos *procesales* que solo pueden ser practicados por la parte, independientemente de la intervención de su abogado. Este elemento tiene repercusión, por ejemplo, en la forma de comunicación para la práctica del acto procesal —se exige, como regla general, la *notificación*

personal (art. 242 del CPC brasileño) o la *comunicación personal a la parte* para la validez de la aplicación de consecuencias por la inobservancia del comportamiento esperado— y en la propia especie de sanción aplicable ante la inobservancia de ese comportamiento.[21]

Además, esta práctica *personal* del comportamiento es fundamental para fines instructivos, en la medida en que la prueba obtenida *por medio de una determinada persona* presupone que esa persona, específicamente, consigue colaborar *con su conocimiento o sus sentidos* para la determinación de los hechos relevantes para el proceso. Por ello, la ley procesal habla de «comparecencia personal de las partes» para el interrogatorio libre (art. 139, VIII del CPC brasileño) y de intimación personal para la recolección del testimonio de las partes (art. 385, § 1º del CPC brasileño), a fin de permitir que la parte responda personalmente (art. 387 del CPC brasileño) sobre los hechos relevantes.

Si la obtención de ciertas pruebas depende del conocimiento específico o de los sentidos de alguien, sería impensable admitir que otro pudiera estar en su lugar para este fin.[22]

En todo caso, dado que el objetivo de este texto es apreciar el comportamiento practicado por las *partes* y su uso como argumento de prueba en el proceso, los actos practicados por medio del abogado deben quedar fuera de este análisis. Aunque el tema es sumamente interesante, exige un enfoque distinto y análisis que desbordan los límites de este pequeño ensayo. Al fin y al cabo, si el comportamiento es adoptado por el abogado, entonces sería adecuado pensar en qué consecuencias debe aportar esta conducta *para él*, y no para la parte. Transferir a la parte las consecuencias de un comportamiento realizado por su abogado —aunque sea un pensamiento muy común en las legislaciones procesales[23]— es suponer, como hacen las legislaciones

---

[21] Así, por ejemplo, el art. 528 del CPC brasileño, exige la intimación personal del deudor para la ejecución de prestación alimenticia, sobre todo porque, solo con esa comunicación personal es posible, después, imponer la grave sanción de la prisión civil, en caso de no pago de la prestación debida (art. 528, § 3º del CPC brasileño).

[22] De cierto modo, es por esa razón que el CPC brasileño, aunque admite el testimonio de incapaz, prevé que «la confesión hecha por un representante solamente es eficaz en los límites en que este puede vincular al representado» (art. 392, § 2º del CPC, traducción libre). Al fin y al cabo, no se trata ahí solo de un problema de determinación de hechos. Más que eso, se trata de vincular al representado —que efectivamente tiene el conocimiento de los hechos— al establecimiento de una verdad sobre cierto hecho que se vuelve indiscutible. Si solo la parte —que presenció los hechos— puede atestar cómo esos hechos ocurrieron, un tercero solo puede manifestar esa declaración de conocimiento en los límites en que se le permite ese tipo de manifestación. Además, bajo otro aspecto, reglas así también tienen importante función «negocial». Si esa confesión puede implicar la dispensa de prueba sobre cierto hecho, se tiene ahí una convención procesal entre las partes, que solo puede ser practicada por ella misma o por quien tenga poderes para vincularla en actos negociales.

[23] La ley procesal brasileña, por ejemplo, hace constantemente esa asimilación. Considera que la parte es litigante de mala fe cuando *su abogado* (ya que es él quien practica los actos procesales por regla) deduce pretensión o defensa infundadas, provoca incidentes infundados o interpone recurso de contenido dilatorio (art. 80 del CPC). Ya en relación al abogado, dice el Código de Proceso Civil, su

procesales, que existe una verdadera simbiosis entre abogado y parte, que no siempre refleja lo que ocurre en la realidad. Y, si esa simbiosis es comúnmente aceptada para fines procesales generales, parece que es excesiva en lo que se refiere al campo de la prueba. Tomar, *a partir del comportamiento del abogado de la parte*, argumentos de prueba en perjuicio de la parte —sin ningún cuestionamiento ulterior sobre la alineación entre esa conducta y el interés de la parte representada— suena exagerado y erróneo, aunque, como se verá más adelante, también existan ejemplos en los que la ley concluye abiertamente de esa manera.[24] Al fin y al cabo, en muchas circunstancias son diversos los intereses del abogado y de la parte representada, no pudiéndose excluir aún otros innumerables factores que influyen en la relación abogado-cliente, lo que hace injusta la inferencia, únicamente a partir del comportamiento del abogado y sin mayores cautelas, de pruebas contra la parte.[25]

Todas estas circunstancias, en fin, recomiendan dejar el análisis de estas interrelaciones para otro espacio, concentrando aquí el foco en la evaluación del comportamiento de la parte y su impacto para fines instructivos.

c) Las conductas que constituyen objeto de la deliberación judicial (causa de pedir y pedir) y las conductas que pueden comprobar los hechos insertos en la causa de pedir y en la defensa. Otra distinción que puede ser relevante en este campo se refiere a la posición de la conducta en el proceso. Es posible que las conductas de las partes constituyan el verdadero objeto de la decisión judicial. El mérito de la demanda puede consistir en saber si un determinado comportamiento es lícito o no, o si cierta conducta es válida/eficaz o no. El juez puede ser llamado a decidir si expresar un concepto negativo respecto a otra persona constituye una manifestación de la libertad de opinión o si, por el contrario, implica una violación de la imagen de esta persona. Aquí, el objeto de la decisión judicial implica precisamente la valoración del comportamiento del sujeto para la aplicación de la consecuencia jurídica deseada por la parte actora. La conducta ahí constituye la *causa petendi* de la demanda judicial y es el punto neurálgico de la decisión judicial. Una situación idéntica es aquella en la que se pretende evaluar si, por conducir alcoholizado un vehículo automotor e involucrarse en un accidente automovilístico, deben considerarse presentes los requisitos del art. 186 del

---

eventual responsabilidad es solo *disciplinaria* y debe ser investigada «por el respectivo órgano de clase o corregiduría» (art. 77, § 6º del CPC), no pudiendo ser objeto de sanción por el juzgado de la causa.

[24] Es lo que ocurre con la llamada tutela de la evidencia, en que la ley brasileña autoriza la satisfacción provisional de la pretensión inicial cuando el juez entienda que hay abuso del derecho de defensa o el evidente propósito procrastinador [*protelatório*] de la parte (art. 311 del CPC). Aquí, aunque el comportamiento ilícito sea, por regla, practicado por el abogado, el código acaba por concluir que esa conducta (del abogado) es indicativo de que la parte demandada no tiene razón en la resistencia que ofrece a la pretensión del autor. Y, por eso, permite al autor que disfrute, provisionalmente, del bien hasta la solución final de la causa. Como también se verá, hay controversia sobre esa conclusión de que la tutela de la evidencia implica presunción judicial derivada de la conducta de la parte. Pero esa ha sido la lectura más usual en el derecho brasileño, razón por la cual el caso es empleado como ejemplo en este texto.

[25] Sobre un debate más amplio al respecto de las relaciones entre abogado-cliente, v., Arenhart (2025, p. 262 y ss.).

Código Civil brasileño y, consecuentemente, si incide la responsabilidad civil por los perjuicios causados (art. 927 del Código Civil brasileño). Los comportamientos de los sujetos, en estos casos, son el soporte material de la incidencia de la regla jurídica cuya consecuencia se pretende aplicar. La conducta de conducir alcoholizado y causar un accidente es (o puede ser) calificada como un acto ilícito que impone el deber de reparar los daños y, en esa dimensión, esa conducta es el verdadero *hecho jurídico* del cual deriva la consecuencia prevista por el ordenamiento jurídico.

Idéntico razonamiento podría aplicarse a las defensas eventualmente presentables por el demandado. Cuando el demandado alega la excepción de contrato no cumplido (*exceptio non adimpleti contractus*), por ejemplo, afirma una defensa que se sostiene en la inexistencia de un comportamiento por parte del actor que sería requisito indispensable para el ejercicio de la pretensión deducida en el proceso. Aquí, nuevamente, la conducta (ahora del actor) es objeto de la deliberación judicial y su evaluación es fundamental para el análisis sobre la procedencia o improcedencia de la demanda inicialmente formulada.

En esta misma línea, sería posible pensar la conducta como la propia «consecuencia» establecida por la norma. Así, por ejemplo, la conducta de «reparar el daño» —que sería la consecuencia de ambos ejemplos dados anteriormente—, o de restituir cierta cosa (v.g., art. 952 del Código Civil brasileño), o incluso de abstenerse de violar la privacidad de alguien (art. 21 del Código Civil brasileño), puede ser considerada como *el pedido* vehiculado en cierta demanda judicial. Y, en esos casos, la eventual «prueba de la conducta» puede ser un elemento a considerar como defensa material indirecta (v.g., el pago, el cumplimiento de la prestación o la imposibilidad de este pedido), que también se inserta en el fondo del proceso.

Diferente es la situación en que la conducta de la parte es empleada para *demostrar otro* hecho que constituye el antecedente fáctico de la norma o que pueda contribuir a demostrar la satisfacción del consecuente normativo buscado en juicio.

Tómese la situación hipotética en que alguien pretende demostrar que cierto contrato debe ser anulado porque fue celebrado con vicio del consentimiento, como el dolo (art. 145 del Código Civil brasileño). Para demostrar este vicio de consentimiento, la parte actora argumenta que fue engañada por medio de artilugios empleados por el demandado, quien se aprovechó de la ingenuidad de la parte actora para inducirla a celebrar el negocio. Para demostrar su ingenuidad, la parte actora argumenta que no posee instrucción, que es una persona simple, que vive en el medio rural y, por lo tanto, no está acostumbrada a ese tipo de negocio jurídico. Aquí, el hecho de que esa persona no posea suficiente discernimiento, pudiendo haber sido fácilmente engañada, puede ser probado mediante la demostración de los hábitos que esa persona presenta, por su conducta pretérita e incluso por su comportamiento durante la celebración del negocio jurídico.

En este caso, el comportamiento de la parte es empleado como prueba de un elemento intencional (subjetivo) que difícilmente podría ser demostrado por las vías tradicionales.

Lo mismo puede decirse en relación con la llamada omisión intencional (art. 147 del Código Civil brasileño). Según el código civil, este silencio *intencional* puede ser causa de anulación del negocio jurídico, «probándose que sin ella el negocio no se habría celebrado».[26] ¿Cómo sería posible demostrar que este silencio fue «intencional», y no solo derivado del olvido de la otra parte? ¿Y cómo se demostraría que, si se hubiera revelado la circunstancia omitida, el negocio no se habría celebrado?

Aquí, nuevamente, solo las circunstancias que rodean la celebración del negocio pueden indicar estos elementos.[27] Y la demostración de estos elementos, que no constituyen *puntos esenciales de la causa de pedir o de la defensa*, acaban elevándose a caracteres indispensables para matizar los hechos «principales», permitiendo su adecuada caracterización.

Sin duda, ambas perspectivas son desafiantes y ofrecen enfoques importantes que realzan el examen del comportamiento de los sujetos para el proceso. Sin embargo, el primer aspecto se relaciona más con alguna «prueba de la conducta de las partes», mientras que el segundo sí se alinea con la idea del empleo de la «conducta de las partes como prueba». Por ello, y nuevamente dadas las limitaciones de esta investigación, se elige esta segunda categoría como objetivo de este texto.

d) La conducta aislada y la conducta reiterada. La elección de la categoría antes mencionada no excluye, por otro lado, que se puedan agregar otros elementos para el análisis de este tipo de conducta. Así, también se podrían clasificar las conductas de las partes por medio de su frecuencia, lo que puede contribuir incluso a la propia fuerza probatoria del comportamiento pasado.

Así, se pueden separar las *conductas aisladas* de las *conductas reiteradas.* Es cierto que aquí, la clasificación puede ligarse a aspectos subjetivos de la evaluación de la conducta y a su potencialidad para servir como elemento de prueba.

La reiteración de cierto comportamiento puede eventualmente servir como indicativo del carácter o de la inclinación de comportamiento de alguien,[28] lo que quizás no sería posible evaluando una única conducta aislada. Además, aunque esta

---

[26] Dice la regla del Código Civil brasileño: «art. 147. Nos negócios jurídicos bilaterais, o silêncio intencional de uma das partes a respeito de fato ou qualidade que a outra parte haja ignorado, constitui omissão dolosa, provando-se que sem ela o negócio não se teria celebrado.»

[27] Atento a la dificultad probatoria en casos así, establecía el Código de Proceso Civil brasileño de 1939, ya revocado, que «el dolo, el fraude, la simulación y, en general, los actos de mala fe podrán ser probados por indicios y circunstancias» (art. 252). Además, añadía que «en la apreciación de los indicios, el juez considerará libremente la naturaleza del negocio, la reputación de los indiciados y la verosimilitud de los hechos alegados en la inicial y en la defensa» (art. 253).

[28] La afirmación no deja de ser polémica, en la medida en que varios países prohíben el empleo de la llamada «*character evidence*», particularmente para fines de acusación criminal, exactamente porque es posible que se le dé un valor excesivo, aflorando prejuicios del jurado en relación con ciertas personas. No obstante, se ha defendido frecuentemente también que, en casos en que el análisis del carácter de la persona es esencial para la solución de la controversia, debe admitirse la prueba de ese carácter, lo que puede hacerse por medio de cualquier medio, incluso por el análisis de la conducta pretérita del sujeto

reproducción pueda llevar a algún indicativo con relación al carácter de alguien o a la inclinación de sus hábitos, estos indicativos en regla son muy débiles para conclusiones determinantes respecto a la prueba de ciertos hechos.[29]

Por otro lado, como se verá, hay casos en que un único comportamiento sirvió de base para una determinada conclusión, a pesar de no haber ninguna evidencia de que aquella conducta sería reproducida en el futuro o fue reproducida en otros momentos.

Los hábitos, así, constituyen un importante punto a debatir en este texto. Como se pretende demostrar más adelante, la verificación de cierto hábito puede sí, en ciertas circunstancias, ser tomada como prueba de algún hecho relevante para el proceso. No obstante, aunque la reiteración de la conducta pueda reforzar la justificación —haciéndola más fácil o dando mayor credibilidad a la justificación dada por la decisión—, parece que todavía hay espacio para que situaciones y conductas puntuales puedan también tener un lugar asegurado en este ambiente.

e) La conducta comisiva y la omisión. Por otra parte, y más como una aclaración, no se puede olvidar que, para fines jurídicos, la omisión puede muchas veces ser interpretada también como una acción (como una acción negativa). Por eso, cuando se piensa en la prueba por la conducta de las partes, es posible que la omisión de estos sujetos también sea interpretada como argumento de prueba.

Diversas veces, por cierto, es exactamente lo que hace la legislación procesal. En numerosos casos, las leyes procesales —como se verá a continuación— atribuyen a la ausencia de comportamiento de alguien efectos semejantes a ciertas conductas activas o, en otras ocasiones, vinculan a esa omisión ciertas presunciones que dispensarán pruebas de algunos hechos en el futuro. De este modo, para los fines aquí discutidos, se tratarán como asimiladas las nociones de conducta positiva y de omisiones, salvo ante previsiones legales expresas, que impongan otra interpretación.

f) Hecho concreto, categoría de hechos y la conducta de los sujetos. Finalmente, hay otra modalidad de hechos que también se mencionará —aunque brevemente— en este texto. Por regla general, los hechos comprendidos y tratados por el proceso involucran hechos determinados y aislados, ocurridos en el mundo sensible. Es respecto a estos hechos que se hacen afirmaciones que serán, posteriormente, comprobadas en juicio.

No obstante, es posible que el proceso judicial también reciba lo que muchos autores denominan «hechos legislativos»,[30] es decir, hechos que pueden informar a

---

(v., por ejemplo, *Rule 404* de las *Federal Rules of Evidence* y, en la jurisprudencia civil norteamericana, *Pino v. Koelber*, 389 So. 2d 1191, 1193 (Fla. Dist. Ct. App 1980)).

[29] Como señala Schauer, «these days most psychologists recognize that explaining behavior is not only a matter of understanding dispositions or traits, and not only a matter of understanding the situation. Instead, human behavior is a function of the *interaction* between dispositions and situations. Different people react to the same situations in different ways, and thus it is a mistake to think that it is all about the situation» (Schauer, 2022, p. 214).

[30] La expresión fue primeramente empleada por Kenneth Culp Davis (v., Davis, 1955, *passim*).

la Corte sobre ciertas decisiones legislativas en materia de Derecho y Política (Davis, 1942, p. 404). De modo más amplio, podría decirse que aquí se pretende tratar de «hechos abstractos», considerados de modo genérico como presupuesto para la incidencia de ciertas normas. En contraste, los hechos regularmente traídos al proceso serían aquellos denominados «hechos adjudicativos» (*adjudicative facts*), consistentes en saber lo que las partes hicieron, cuáles fueron las circunstancias en que algo ocurrió o cuáles serían las condiciones antecedentes de alguna cosa (Davis, 1942, p. 402). Los primeros se refieren a juicios realizados por los tribunales en relación con la interpretación de textos normativos, mientras que los hechos adjudicativos son relevantes para el análisis de algún caso en particular (Faigman, 2008, p. 44).

Así, saber si A estuvo en determinado lugar en cierto momento y si estaba alcoholizado al conducir el vehículo que se vio involucrado en un accidente corresponde a un hecho adjudicativo. En cambio, la circunstancia de que conducir alcoholizado reduce los reflejos del conductor es un «hecho abstracto», que puede ser reconducido a la noción de hecho legislativo o, en términos más generales, a la idea de una *categoría de hechos*.[31]

Como regla general, la determinación de los hechos concretos depende de pruebas ligadas al conocimiento *específico* de la ocurrencia circunstancial de ese hecho. En relación con los hechos abstractos, las pruebas normalmente se relacionan antes con la aceptación general de la validez de cierta inferencia (lo que, normalmente, se relaciona con el conocimiento técnico-teórico o con las máximas de la experiencia) que con otros elementos.

También esta clasificación será útil en este texto, porque, como se verá, la determinación de categorías de hechos puede deducirse más fácilmente de comportamientos específicos de las partes que la especificación de hechos puntuales.

Finalmente, es importante subrayar que, a veces, es difícil establecer la diferencia entre la prueba del hecho (relevante para la solución de la controversia) y la prueba de la conducta que comprobará ese hecho. En la medida en que, al fin y al cabo, todos estos elementos servirán para la demostración del hecho relevante para el proceso, puede haber alguna dificultad en la separación entre los casos en que la conducta es un elemento de la norma a aplicar —y, entonces, su prueba constituye prueba de uno de los elementos de la hipótesis constante de la norma— y las otras situaciones en que la conducta es empleada como elemento que permitirá considerar existente o no el hecho cualificado por la norma a incidir.

---

[31] En términos exactos, la asimilación entre hechos legislativos y hechos genéricos o categorías de hechos no es precisa. Como aclara Ann Woolhandler, «the key difference between adjudicative and legislative facts is not the characteristics of particular versus general facts, but rather, evidence whose proof has a more established place and a more predictable effect within a framework of established legal rules as distinct from evidence that is more manifestly designed to create the rules» (Woolhandler, 1988, p. 114). Aun así, los hechos legislativos no dejan de ser también hechos generales, que pueden ser calificados dentro de esa noción más amplia de «categoría de hechos».

Lógicamente, sin embargo, sería imposible, en los estrechos límites de un artículo, pretender analizar todos los casos en que alguna conducta es presupuesto para la incidencia de la norma y, en esos casos, su prueba. Este examen, además, parece estéril porque llevará a la inexorable conclusión de que cualquier prueba es hábil para la demostración de esa conducta.

Por lo tanto, en este texto se opta por apreciar solo las situaciones en que la conducta es evaluada como elemento que puede demostrar los hechos a probar, es decir, los hechos litigiosos que conforman la norma jurídica cuya incidencia (o no incidencia) se pretende en el proceso.

Analizando, por ejemplo, la regla del art. 187 del Código Civil de Brasil – que prevé que «también comete acto ilícito el titular de un derecho que, al ejercerlo, excede manifiestamente los límites impuestos por su fin económico o social, por la buena fe o por las buenas costumbres» –, la prueba del ejercicio del derecho implica prueba de conducta que constituye un elemento de la norma mencionada. Por otro lado, la prueba de conductas que puedan indicar los límites de la «buena fe», del «fin económico o social» o la dimensión de las «buenas costumbres» constituye un ejemplo de aquello que aquí se espera examinar. Es un caso en que el comportamiento de los sujetos puede llevar a la *interpretación* de que cierto hecho relevante para el proceso realmente ocurrió.

Como se ve, para la interpretación de este hecho, es posible considerar, eventualmente, un hecho aislado, conductas reiteradas, categorías de hechos, acciones y omisiones. En fin, es posible que la justificación respecto a la ocurrencia del hecho a probar se valga de argumentos fundados en otros comportamientos, que pueden ser suficientes para demostrar la convicción necesaria para tener como ocurrido aquel primer hecho.

Es dentro de estos límites que se pretende aquí desarrollar el asunto.

3. EFECTOS *LEGALMENTE* PREVISTOS PARA CIERTAS CONDUCTAS PROCESALES Y CLÁUSULAS GENERALES DE INTERPRETACIÓN DE ESAS CONDUCTAS

En la gran mayoría de los sistemas procesales, se atribuye a ciertos comportamientos de las partes una significativa importancia probatoria expresa. En muchos casos, algunos comportamientos procesales son interpretados, por reglas específicas, como determinantes para la formación de la convicción de la autoridad judicial.

Tal ocurre con el fenómeno de la rebeldía. Aunque esté fuera del ámbito de estudio de este texto, el comportamiento omisivo de la parte (por regla, por intermedio de su abogado) al no presentar defensa en el proceso puede repercutir en la incidencia de una presunción *general* de aceptación de los hechos tal como los alegó

el demandante.[32] También en Brasil, el comportamiento (personal) de las partes en ciertos momentos del proceso puede constituir un importante medio de prueba. Así, la parte que, sin justificación, deja de comparecer al acto designado para la toma de su declaración, o, compareciendo, se niega a declarar, o, aún, deja de responder a lo que se le pregunta o emplea evasivas en sus respuestas, puede tener aplicada en su contra la consecuencia de la «confesión ficta», presumiéndose verdaderos los hechos que con el testimonio personal se pretendía demostrar.[33] De modo semejante, el art. 400 del CPC brasileño prevé que, solicitada la exhibición de un documento o cosa contra la parte, el juez debe presumir como verdadero el hecho que se quería probar con esa exhibición si la parte no ha presentado el documento o la cosa, o si su negativa a exhibirlo es considerada ilegítima.

Además de estas reglas más amplias, la legislación contempla diversas presunciones puntuales que toman en consideración el comportamiento de las partes para establecer pruebas en uno u otro sentido. En esta línea, el art. 232 del Código Civil brasileño afirma que «la negativa a la pericia médica ordenada por el juez podrá suplir la prueba que se pretendía obtener con el examen». También las reglas sobre imputación del pago (arts. 352 a 355 del Código Civil brasileño) expresan argumentos de prueba tomados a partir del comportamiento del acreedor y del deudor en la satisfacción de múltiples deudas de la misma naturaleza.

También otros países tienen preceptos similares. Es el caso de Italia que también admite la confesión ficta, tomada a partir de la no presentación de la parte o de su negativa a declarar (art. 232 del CPC),[34] así como la presunción de veracidad resultante de la no exhibición de la cosa (art. 210, 4ª parte del CPC). La confesión ficta también está prevista, con variaciones, en sistemas como el español (arts. 304 y 307 de la LEC), el colombiano (art. 205 del Código General del Proceso) y el argentino (art. 417 del CPC de la Nación). Para el derecho portugués, se autoriza al tribunal a considerar «libremente» el valor dado a la conducta de la parte que «no comparezca o se niegue a declarar o a prestar informaciones o aclaraciones, sin probar justo impedimento, o responda que no recuerda o no sabe nada» (art. 357º del CC). Similar es la previsión contenida en el derecho alemán (§§ 446 y 454 del ZPO). Francia, diferentemente, trata la ausencia de la parte o su negativa a declarar como equivalente a un «comienzo de prueba por escrito» (art. 198 del CPC), lo

[32] Es lo que expresamente prevé el art. 334 del CPC brasileño, aunque existan excepciones a la incidencia de esa presunción.

[33] Arts. 385, § 1º y 386, del CPC brasileño.

[34] El derecho italiano va más allá, previendo también la posibilidad de obtener argumentos de prueba a partir de la negativa en el ofrecimiento del juramento decisorio («articolo 239. La parte alla quale il giuramento decisorio è deferito, se non si presenta senza giustificado motivo all'udienza all'uopo fissata, o, comparendo, rifiuta di prestarlo o non lo riferisce all'avversario [234 c.p.c.], soccombe rispetto alla domanda o al punto de fatto relativamente al quale il giuramento è stato ammesso; e del pari soccombe la parte avversaria, se rifiuta di prestare il giuramento que le è riferido. Il giudice istruttore, se ritiene giustificata la mancata comparizione della parte che deve prestare il giuramento, provvede a norma dell'articolo 232 secondo comma»).

que, de todos modos, implica atribuir consecuencias a este comportamiento de la parte. La presunción de veracidad derivada de la no exhibición de documentos o cosas, a su vez, es también común, estando prevista en sistemas como el alemán (§ 427 del ZPO), el español (art. 283-bis, h, n. 1 de la LEC), el portugués (art. 417, n. 2, c/c art. 430 del CPC), el colombiano (art. 267 del CGP) y el argentino (art. 388 del CPC de la Nación).

Figuras como la rebeldía, además de implicar consecuencias probatorias —dispensando, al menos para el sistema brasileño y para otros que le atribuyen consecuencias similares, la prueba de ciertos hechos (art. 374 del CPC brasileño)—, funcionan también como técnica de simplificación procedimental, acelerando la tramitación del proceso. No hay duda, sin embargo, de que, al menos desde el punto de vista probatorio, figuras como esta también permiten extraer argumentos de prueba, sobre todo para sistemas como el brasileño, en el que la presunción de veracidad del hecho derivada de la rebeldía *no es vinculante para el juez*.

Junto a hipótesis más conocidas como la de la rebeldía, los sistemas jurídicos suelen contener innumerables otras reglas con presunciones específicas ligadas a ciertos comportamientos de las partes. Solo para citar el derecho brasileño, es lo que se ve en el art. 1.198, párrafo único («aquel que comenzó a comportarse del modo como prescribe este artículo, en relación con el bien y a la otra persona, se presume detentador, hasta que pruebe lo contrario») o del art. 1.324 («el condómino que administre sin oposición de los otros se presume representante común»), ambos del Código Civil brasileño. De modo similar, se tiene el art. 232 del Código Civil brasileño – que trata de la negativa a someterse a una prueba pericial, permitiendo que el juez «supla», en razón de esa negativa, la prueba que se pretendía obtener con el examen –, previsión que constituye una mera consolidación de un entendimiento jurisprudencial formado desde hace mucho tiempo.[35]

De todas formas, junto a estas previsiones *expresas* y puntuales, parece importante analizar si, actualmente, los sistemas procesales contienen una cláusula *general* que autorice a considerar el comportamiento de las partes como argumento para fines probatorios.

En Brasil, aunque la jurisprudencia admita excepcionalmente el uso de la conducta procesal de las partes para fines probatorios, no existe ninguna regla general del proceso civil que, de forma expresa y general, autorice al juez a extraer argumentos de prueba de este comportamiento. Esta posibilidad se extrae, como se verá con más detalle más adelante, del régimen de atipicidad probatoria, ofrecido por el art. 369 del CPC, combinado con el criterio de «convicción motivada» adoptado por la ley brasileña para la valoración judicial de la prueba (art. 371 del CPC), pero no existe una regulación legal específica sobre el tema.

---

[35] V., Theodoro Jr. (2003, p. 574/578). En la jurisprudencia, este entendimiento fue cristalizado por la *Súmula* 301 del Superior Tribunal de Justicia, que establece que «em ação investigatória, a recusa do suposto pai a submeter-se ao exame de DNA induz presunção *juris tantum* de paternidade».

En algunos otros países, sin embargo, la cuestión recibe tratamiento específico. En Italia, el art. 116 del CPC establece que el juez puede extraer «argumentos de prueba» de las respuestas que las partes den en los interrogatorios libres, de su negativa injustificada a permitir las inspecciones ordenadas y, de modo más amplio, del comportamiento de las partes en el proceso. También en el derecho colombiano existe una previsión expresa en el sentido de que «el juez podrá deducir indicios de la conducta procesal de las partes» (art. 241 del CGP). La legislación argentina autoriza al juez a emplear la conducta de las partes como argumento para «corroborar» su convicción formada a partir de las pruebas.[36] El CPC peruano, de forma más restringida, permite al juez extraer «conclusiones contra los intereses de las partes» a partir de la conducta de estas *en el proceso*, sobre todo cuando dejan de colaborar para la obtención de medios de prueba o actúan de forma que obstruyen la secuencia del proceso.[37] En los Estados Unidos de América, las *Federal Rules of Evidence* autorizan, de forma excepcional y expresa, el empleo de la prueba de la personalidad de las personas para fines criminales, particularmente para la prueba de elementos que indiquen aspectos subjetivos del crimen practicado por el delincuente o para la prueba de datos referentes al comportamiento pasado de la víctima.[38]

---

[36] Dice el art. 163, n. 5, del Código Procesal Civil y Comercial de la Nación que «La conducta observada por las partes durante la sustanciación del proceso podrá constituir un elemento de convicción corroborante de las pruebas, para juzgar la procedencia de las respectivas pretensiones.»

[37] «Artículo 282. El Juez puede extraer conclusiones en contra de los intereses de las partes atendiendo a la conducta que estas asumen en el proceso, particularmente cuando se manifiesta notoriamente en la falta de cooperación para lograr la finalidad de los medios probatorios, o con otras actitudes de obstrucción. Las conclusiones del Juez estarán debidamente fundamentadas.»

[38] «Rule 404. Character Evidence; Other Crimes, Wrongs, or Acts. (a) CHARACTER EVIDENCE. (1) Prohibited Uses. Evidence of a person's character or char- acter trait is not admissible to prove that on a particular occasion the person acted in accordance with the character or trait. (2) Exceptions for a Defendant or Victim in a Criminal Case. The following exceptions apply in a criminal case: (A) a defendant may offer evidence of the defendant's pertinent trait, and if the evidence is admitted, the prosecutor may offer evidence to rebut it; (B) subject to the limitations in Rule 412, a defendant may offer evidence of an alleged victim's pertinent trait, and if the evidence is admitted, the prosecutor may: (i) offer evidence to rebut it; and (ii) offer evidence of the defendant's same trait; and (C) in a homicide case, the prosecutor may offer evidence of the alleged victim's trait of peacefulness to rebut evidence that the victim was the first aggressor. (3) Exceptions for a Witness. Evidence of a witness's Character may be admitted under Rules 607, 608, and 609. (b) OTHER CRIMES, WRONGS, OR ACTS. (1) Prohibited Uses. Evidence of any other crime, wrong, or act is not admissible to prove a person's character in order to show that on a particular occasion the person acted in accordance with the character. (2) Permitted Uses. This evidence may be admissible for another purpose, such as proving motive, opportunity, intent, preparation, plan, knowledge, identity, absence of mistake, or lack of accident. (3) Notice in a Criminal Case. In a criminal case, the prosecutor must: (A) provide reasonable notice of any such evidence that the prosecutor intends to offer at trial, so that the defendant has a fair opportunity to meet it; (B) articulate in the notice the permitted purpose for which the prosecutor intends to offer the evidence and the reasoning that supports the purpose; and (C) do so in writing before trial —or in any form during trial if the court, for good cause, excuses lack of pretrial notice.»

4. EL DERECHO BRASILEÑO Y LA INTERPRETACIÓN GENERAL DE LA CONDUCTA DE LAS PARTES PARA FINES PROBATORIOS

Ya se ha visto que el ordenamiento jurídico brasileño posee reglas puntuales que disciplinan las consecuencias de ciertas conductas de las partes en el proceso. Sin embargo, no existe una regla general que aborde el asunto. De ahí surge el primer desafío al abordar el tema: saber si, a pesar de la falta de una previsión general, el comportamiento de los sujetos procesales puede ser interpretado como argumento de prueba.

La conducta de las partes puede, sin duda, ser calificada de forma amplia como un hecho jurídico. Si es cierto, al final, que ese comportamiento es jurídicamente relevante y puede ser determinante en el desarrollo del proceso y en los caminos que el proceso habrá de seguir, entonces no hay duda de que esas conductas —e, incluso, la inacción de la parte— conducen a consecuencias jurídicas *dentro del proceso* (Righi, 1981, p. 2), y, así, pueden ser tomadas como hechos jurídicos procesales. Más precisamente, se trata de un *acto jurídico procesal*, en la clasificación adoptada por Carnelutti (1944, p. 69), en la medida en que se trata de un acto en que la voluntad humana tiene relevancia específica para la calificación de ese hecho.

Como regla general, los actos jurídicos generan, como consecuencias, aquellas descritas de forma expresa en la ley. Sin embargo, como ya se ha señalado, el sistema brasileño, en la línea de lo que también hacen varios otros países, no posee una regulación específica sobre las consecuencias *probatorias* de la conducta de las partes. Así, cabría preguntarse cómo debe interpretarse este comportamiento, específicamente en lo que respecta a la valoración probatoria de esa conducta.

El sistema brasileño adopta, como regla general de valoración de las pruebas, el criterio de la *convicción motivada*, de modo que, en los términos del art. 371 del CPC, «el juez apreciará la prueba constante en los autos, independientemente del sujeto que la haya promovido, e indicará en la decisión las razones de la formación de su convencimiento». En un sistema así, se considera que la inexistencia de una regla específica que autorice a valorar la conducta de las partes para fines probatorios es irrelevante. En consecuencia, aunque no exista una regla general que lo autorice, se entiende que el juez sí puede valerse de esa conducta para formar su convicción (Rigi, 1981, p. 3/4).

Según este modelo, solo se exige que el comportamiento de la parte esté demostrado en los autos. Así, el juez no puede considerar la conducta de la parte conforme a lo que sea de su conocimiento privado o según rumores que haya oído de forma extrajudicial. Es necesario que haya, en los autos, la demostración de cierto comportamiento —único o reiterado— de la parte y esto autorizará que esa conducta sea evaluada para fines instructivos.

Por otro lado, en el sistema brasileño —en la línea de lo que también es adoptado en la amplia mayoría de los países—, se adopta el principio de la atipicidad de las pruebas. Por lo tanto, aunque no esté regulada de forma expresa, la conducta de las

partes puede ser tomada por el juez, en la medida en que, conforme a lo previsto en el art. 369 del CPC, «Las partes tienen el derecho de emplear todos los medios legales, así como los moralmente legítimos, aunque no especificados en este Código, para probar la verdad de los hechos en que se funda el pedido o la defensa e influir eficazmente en la convicción del juez».

Concluir que la conducta de las partes puede ser tomada para fines probatorios, sin embargo, aún es insuficiente. Es necesario saber a qué régimen debe someterse este elemento de prueba.

Por regla general, en Brasil, se entiende que esta conducta constituye *prueba indiciaria*, de modo que este comportamiento, alineado con otros indicios similares, puede apuntar a la comprobación de cierto hecho que se espera demostrar. Así, por ejemplo, se entiende como razonable concluir que, demostrada la destrucción de autos o de cierto documento por acto voluntario de una de las partes, había allí prueba en beneficio del adversario. Tal conducta, por lo tanto, ciertamente puede implicar argumentos de prueba contra la parte que así actuó.[39]

En esta misma línea, la doctrina nacional concluye que sería razonable suponer que, en determinados casos, el empleo de medidas dilatorias en el proceso fuera interpretado como prueba contra la parte, aunque no haya previsión legal al respecto.[40]

---

[39] El ejemplo es de Ivan Righi, en el texto mencionado.

[40] Para gran parte de la doctrina nacional, constituye una derivación de esa idea de que el comportamiento abusivo de la parte implica una señal de ausencia de defensa seria —aunque exista divergencia doctrinaria sobre la interpretación del alcance de las expresiones empleadas por el CPC brasileño (v., por todos, Marinoni, 2011, p. 56 y ss.; Mitidiero, 2022, p. 165 y ss.)— la llamada «tutela de la evidencia» (especie de tutela provisional, concedida de forma anticipada, autorizando a la parte autora a la fruición del bien que constituye el objeto del proceso hasta la decisión final), que puede ser conferida, entre otras hipótesis, «cuando, quedare caracterizado el abuso del derecho de defensa o el manifiesto propósito procrastinador de la parte» (art. 311, inc. I del CPC brasileño – traducción libre). Para la visión dominante, este comportamiento procrastinador constituiría un indicio de que la parte demandada no tiene razón y, por lo tanto, busca abusar del proceso como forma de postergar la satisfacción del derecho de la parte actora que tiene razón. V.g., Dotti (2020, p. 238) y ss.; Mazini, (2020, p. 49); Lopes (2006, p. 55). Como pondera Rogéria Dotti, «o abuso do direito de defesa ou a prática de ato processual imbuído de manifesto propósito protelatório acaba por *reforçar a plausibilidade do direito*. Isso porque, após a defesa abusiva, os elementos de convicção produzidos pelo autor ganham ainda mais credibilidade. Chega a se dizer que ocorre, nesse caso, um *salto da verossimilhança para a certeza*. É bem verdade que por se tratar de técnica baseada em cognição sumária, a tutela da evidência não trabalha nos domínios da certeza. Mas é interessante perceber como a conduta abusiva reforça a impressão inicial a respeito da probabilidade» (Dotti, 2020, p. 241). Combinándose, así, este indicio (que lleva a la conclusión de la inconsistencia de la defensa) con la apariencia del derecho de la parte actora, puede el juez anticipar la tutela en favor del requirente, hasta la final solución de la controversia. En verdad, para parte de la doctrina, este comportamiento abusivo, en el modelo actual, trasciende los estrictos límites probatorios, convirtiéndose en verdadero requisito autónomo para el otorgamiento de la medida provisional. En otras palabras, hoy, muchos consideran que ni siquiera se debe cogitar en qué medida este comportamiento implica o no el reconocimiento de la procedencia de la petición inicial. Verificada la ocurrencia de conducta considerada como dilatoria, está satisfecho el requisito que, con la verosimilitud de la alegación de la parte actora, impone al juez el otorgamiento de la medida anticipatoria.

En todos estos casos, la doctrina dominante concluye que el comportamiento de la parte —en el proceso, normalmente— puede ser interpretado como argumento de prueba, asimilado a la noción de prueba indiciaria,[41] lo que, sumado a otros elementos, puede llevar al juez a concluir la ocurrencia o no de determinado hecho relevante para el proceso. El comportamiento de las partes es tomado como un dato probado que, a partir de un razonamiento inductivo o deductivo del magistrado, autoriza a concluir la existencia o no de hechos relevantes para el proceso.

Es cierto que la calificación de esta fuente de prueba como *prueba indiciaria* puede llevar a la creencia de que su valor debe ser *inferior* a los otros medios de prueba. El sistema brasileño, en efecto, no posee ninguna regla que trate del valor probatorio de los indicios de modo general.[42] El Código Civil brasileño de 2002, en su redacción original, poseía una regla que afirmaba que solo se admitiría la prueba indiciaria en los casos en que también fuera admitida la prueba exclusivamente testimonial. El precepto, que vehiculaba un nítido prejuicio contra la prueba indiciaria, sin ningún apoyo técnico, sin embargo, fue expresamente revocado por el Código de Proceso Civil de 2015, de modo que, en el sistema procesal *civil* brasileño en vigor, no hay ninguna regla que trate el asunto de modo exhaustivo.

Este punto será retomado en el siguiente apartado, porque reflexionar sobre la naturaleza de esta prueba parece ser crucial para comprender su alcance.

Sin embargo, por ahora, siguiendo el discurso presentado por la doctrina tradicional, si ese comportamiento es un *indicio*, se concluye que es fundamental evaluar la calidad de ese elemento indiciario (el hecho instructivo) y la del razonamiento empleado por el juez para tener como demostrados los hechos relevantes para el proceso (el hecho a probar).

5. LA CALIDAD DE LA PRUEBA DERIVADA DEL COMPORTAMIENTO DE LAS PARTES

De inmediato, surgen dos cuestiones principales para el análisis de la calidad del indicio aquí tratado y del razonamiento judicial realizado. Por una parte, se impone la pregunta sobre si es o no posible que *solo este indicio, derivado del comportamiento de las partes, sirva para formar la convicción judicial*.

---

[41] En ese sentido, v., Alberto (2005, p. 129); Heñin (2009, p. 75); Cambi y Hoffmann, (2011, p. 84-86); Ferreira (2014, p. 79).

[42] La única regla existente en el sistema nacional sobre la prueba indiciaria es el art. 239 del Código de Proceso Penal brasileño, que se limita a definir ese medio de prueba, afirmando que «se considera indicio la circunstancia conocida y probada, que, teniendo relación con el hecho, autorice, por inducción, a concluir la existencia de otra u otras circunstancias» (traducción libre).

Algunos sistemas normativos, desde este ángulo, sujetan expresamente el uso de las presunciones judiciales[43] (pruebas indiciarias) a la existencia de múltiples indicios. Más allá de eso, se exige por regla que estos indicios sean fuertes y *concordantes* o, en otras palabras, que estos varios indicios converjan hacia el mismo hecho a probar. Así es, por ejemplo, la solución del Código Civil francés, que, además de solo permitir el empleo de la prueba indiciaria cuando la ley permita la prueba del hecho por cualquier medio autorizado por el Derecho, impone que los indicios sean «graves, precisos y concordantes».[44] También es lo que consigna la legislación colombiana, que impone al juez que aprecie los indicios en su *conjunto*, evaluando su «gravedad, concordancia y convergencia», además de su relación con las otras pruebas del proceso.[45] Similar es aún la regla prevista por el Código Procesal Civil y Comercial de la Nación Argentina (art. 163, n. 5, primera parte).

En sentido contrario, sin embargo, otras legislaciones se inclinan por admitir que un único indicio pueda, bajo ciertas circunstancias, llevar a la certeza de la ocurrencia del hecho a probar. Esto ocurre, por ejemplo, con la Ley de Enjuiciamiento Civil española, que afirma que «a partir de un hecho admitido o probado, el tribunal podrá presumir la certeza, a los efectos del proceso, de otro hecho, si entre el admitido o demostrado y el presunto existe un enlace preciso y directo según las reglas del criterio humano» (art. 386, n. 1).[46] De modo más expreso, el Código de Procedimiento Civil Chileno afirma que «una sola presunción puede constituir plena prueba cuando, a juicio del tribunal, tenga caracteres de gravedad y precisión suficientes para formar su convencimiento» (art. 426, segunda parte).

Como ya se mencionó, el derecho brasileño no posee una regla específica que trate de la prueba indiciaria. No obstante, se acepta de modo generalizado la idea de que la conjugación de varios indicios puede conducir a la aceptación de la existencia de cierto hecho. Por regla general, la doctrina y la jurisprudencia brasileñas consideran que un único indicio es demasiado poco para autorizar una conclusión segura sobre la existencia o no de un hecho. Así, volviendo al tema de la conducta de la parte, consi-

---

[43] Jordi Ferrer Beltrán critica, con razón, la expresión presunción judicial, en la medida en que el término puede oscurecer que, en verdad, el razonamiento aplicable para viabilizar el uso de la prueba indiciaria no es distinto del empleado para el razonamiento probatorio en general (v., Ferrer Beltrán, 2022a, p. 425). La expresión, sin embargo, es comúnmente empleada en el derecho brasileño y, por esa única razón, fue utilizada también en el texto.

[44] El art. 1382, del Código Civil francés afirma que «Les présomptions qui ne sont pas établies par la loi, sont laissées à l'appréciation du juge, qui ne doit les admettre que si elles sont graves, précises et concordantes, et dans les cas seulement où la loi admet la preuve par tout moyen».

[45] Dice el art. 242 del Código General del Proceso de Colombia que «El juez apreciará los indicios en conjunto, teniendo en consideración su gravedad, concordancia y convergencia, y su relación con las demás pruebas que obren en el proceso».

[46] La misma regla, dígase, exige que el juez exponga en la sentencia el razonamiento empleado para la presunción y autoriza, en cualquier caso, que la parte perjudicada pueda hacer prueba contraria –tanto en el sentido de demostrar que el hecho a probar no existe, como para indicar que el razonamiento realizado no tiene consistencia (art. 386, n. 2, de la LEC).

derar solo ese comportamiento podría ser algo defectuoso, en la medida en que podría significar varias cosas distintas, no habiendo seguridad de que de él resultaría, inequívocamente, el hecho a ser demostrado.[47] De este modo, por ejemplo, el simple hecho de que alguien haya entregado las llaves de un inmueble a su propietario no sería suficiente para concluir que hay intención de extinguir el contrato de arrendamiento. Sin embargo, sumando ese comportamiento con otros elementos (indiciarios o probatorios) presentados, aumenta el peso de esa fuente de prueba y el grado de impacto que puede tener en la justificación de la decisión judicial. En esta línea, además de la entrega de las llaves, sería importante evaluar el contexto y otras circunstancias, a fin de saber si existía, de hecho, la intención de poner fin al contrato de arrendamiento.

Solo para ilustrar la orientación de la jurisprudencia brasileña —en el sentido de la necesidad de conjugar varios indicios para dar por cierto algún hecho—, tómese el conocido caso de las «píldoras de harina» en Brasil. En 1998, la empresa farmacéutica Schering do Brasil adquirió una máquina empaquetadora para sus medicamentos. Con la intención de probar esta máquina, produjo un lote de un medicamento anticonceptivo, llamado Microvlar, que no poseía ningún compuesto activo, sino solo harina. Por algún problema, sin embargo, estos productos terminaron siendo comercializados, lo que generó posteriormente una avalancha de demandas judiciales, en las que mujeres que decían haber utilizado ese anticonceptivo y haber quedado embarazadas, culpaban a la empresa por el uso del fármaco sin sustancia activa.

Muchas de estas acciones judiciales fueron desestimadas, bajo el entendimiento de que no había prueba suficiente que justificara el reconocimiento del nexo de causalidad y, así, la responsabilización de la empresa. En muchos de estos casos, la conclusión se basó en el hecho de que, a pesar de demostrarse la existencia de receta médica (que indicaría el uso de Microvlar) y la conducta pretérita (el uso de la medicación en diversas otras oportunidades), habría otros elementos que excluirían la plausibilidad del nexo de causalidad.[48] En tantos otros casos, sin embargo, se consideró que la notoria existencia del extravío de las «píldoras de harina», sumada a la prueba del uso regular del medicamento Microvlar, sería suficiente para concluir que el embarazo había sido consecuencia de la ingestión de las píldoras falsas.[49] En otras

[47] Es esta la conclusión de Paulo Osternack do Amaral, al afirmar que, «como regla, esas conductas no podrán ser consideradas aisladamente, sino, analizadas en conjunto con los demás elementos probatorios, podrán subsidiar la conclusión acerca de la ocurrencia o no de los hechos alegados» (Amaral, 2021, p. 121).

[48] Así, por ejemplo, consignó el Superior Tribunal de Justicia en un caso determinado que los indicios derivados del uso de la medicación anteriormente y la existencia de receta del fármaco serían elementos insuficientes porque, en aquel caso específico, porque: a) los envases del producto consumido presentaban fecha de caducidad, lo que no ocurría en las píldoras de prueba; y b) no había indicios de distribución del medicamento sin la sustancia activa en el Estado respectivo en el período en que habría ocurrido el embarazo «ocurrió al final del referido mes, sería improbable que la autora hubiera tenido acceso al dicho lote imperfecto» (STJ, 4ª Turma. REsp 697.078/BA. Rel. Desembargador Convocado Honildo Amaral de Mello Castro. DJe 19.10.09).

[49] Dijo el mismo Superior Tribunal de Justicia en otro fallo que «Com efeito, consta do acórdão recorrido que: (i) 'a apelante estava utilizando o contraceptivo de fabricação da apelada quando engravi-

palabras, la suma de indicios facilitó la justificación del razonamiento judicial, ya sea para considerar probado o para considerar no probado el hecho que sería determinante para la fijación de la responsabilidad en casos muy similares.

Por lo tanto, lo que se observa es que la jurisprudencia nacional concluye habitualmente que indicios aislados no constituyen prueba segura para la formación de juicios sobre hechos. Solo su conjugación y suma con otros elementos serían suficientes para este fin.

Sin embargo, como ha venido resaltando la doctrina más reciente, gran parte de estas conclusiones se apoyan en un prejuicio injustificado. Se supone que el indicio debe tener *menos fuerza* que otros medios de prueba porque en él habría un elemento que no existe en las «pruebas directas», a saber, el razonamiento judicial referente a la presunción, y porque sería incapaz de traducir, de forma inmediata, el hecho a probar, lo que implicaría un menor grado de «certeza» resultante del uso del indicio.[50] Esa es la razón que justificaría la exigencia de la conjugación de varios indicios para dar por probado algún hecho. Esa es también la razón por la cual la jurisprudencia brasileña se resiste, con frecuencia, a permitir que la convicción judicial se apoye únicamente en elementos indiciarios.[51]

Esta suposición, sin embargo, no procede. La supuesta diferencia entre la prueba directa y la prueba indiciaria es absolutamente inútil y no es capaz de presentar ningún elemento de racionalidad que haga a aquella mejor que esta. De hecho, aunque se pueda establecer alguna distinción entre indicio y prueba —involucrando solo el objeto sobre el cual incidirá la prueba—, esta distinción no repercute, por sí sola, en la calidad de la prueba obtenida o en la suficiencia o no de la prueba para la formación del juicio. Como señala Carmen Vázquez (2019), al evaluar los elementos probatorios, el punto relevante se resume en saber si aquello que fue recabado debe o no ser considerado como suficiente para sustentar el convencimiento judicial[52,53]. Y, en esa medida, la diferencia entre indicio y prueba pierde cualquier utilidad.

---

dou, coincidindo a data provável da concepção com o repasse do produto inócuo ao mercado de consumo' (fl. 562, e-STJ); (ii) 'existe um intervalo considerável entre os dois últimos filhos da apelante, razão pela qual supõe-se que soubesse utilizar corretamente o contraceptivo' (fl. 562, e-STJ); e (iii) 'não se pode considerar ter a apelada feito prova robusta sobre a ausência de distribuição do produto defeituoso na região em que vive a apelante. Na realidade, sua prova resume-se no depoimento de uma única testemunha que, por sua vez, não tinha sido formalmente comunicada sobre a irregularidade dos produtos na época em que houve a gravidez indesejada'(fl. 567, e-STJ).» (STJ, 3ª Turma. REsp 1.120.746/SC. Rel. Min. Nancy Andrighi. DJe 24.02.11).

[50] V., por ejemplo, Carnelutti (1947, p. 54 e ss).; Grasso (1973, p. 1306); Montesano y Arieta (1997, p. 94).

[51] V.g., STF, 2ª Turma. HC 77.987. Rel. Min. Marco Aurélio. Dje 10.09.99; STJ, 5ª Turma. AgRg no AREsp 2.791.847/SP. Rel. Min. Joel Ilan Paciornik. Dje 26.05.25.

[52] Como subraya Carmen Vázquez, "Si partimos de la ambigüedad del término prueba, vale la pena distinguir en qué sentido serían diferentes las «pruebas» y los «indicios». Si prueba quiere decir «elemento de juicio» o información a partir de la cual podemos hacer un juicio sobre la verdad o falsedad de la

Nota 53 en página siguiente

Ahora bien, pensar en la *suficiencia* de la prueba implica concluir que, ya sean «pruebas directas» o indicios, lo importante es evaluar la calidad del elemento aportado para la formación de la convicción judicial y, por otra parte, analizar el *conjunto probatorio* y su aptitud para llevar a la convicción de un hecho. Y, siendo así, tanto para pruebas («directas») como para indicios, el conjunto de elementos puede hacer más segura la conclusión sobre el hecho, así como la existencia de un único elemento puede hacer más dudosa su determinación.

Sin embargo, como ocurre con cualquier otro tipo de medio probatorio, la existencia de un único elemento (sea indicio, sea prueba «directa») puede, en muchos casos, ser considerado como suficiente para justificar de modo adecuado la convic-

---

hipótesis, la distinción entre «indicios» y «prueba» es más bien de grado y no categórica. Esto significa que en ambos casos se requiere hacer inferencias y el punto clave radica precisamente en la fuerza de estas. A modo de ejemplo, supongamos que un testigo afirma que ha visto disparar al acusado, el juez tiene que inferir que «el testigo ha disparado al acusado» a partir de la afirmación de aquel; aún más, el testigo efectivamente pudo haber visto disparar al acusado, y el juez en un caso de homicidio tiene que inferir además que la bala disparada fue la que privó de la vida a la víctima (habiendo podido ser el caso de que la víctima muriera de la impresión misma del ataque y el disparo solo hubiera rozado su cuerpo). Los hechos que pueden ser utilizados como premisas de inferencias sobre otros hechos dependerán del conocimiento científico, pero también, y quizá sobre todo, de las máximas de experiencia que utilizan los jueces; y, obviamente, cuáles son las máximas de la experiencia o los conocimientos relevantes dependerá del caso concreto. Por otro lado, si usamos el término «prueba» como resultado probatorio, la diferencia entre pruebas e indicios se ha planteado como si las primeras fueran por sí mismas suficientes para probar los hechos, mientras los segundos no. Sin embargo, el punto problemático no es si se trata de una prueba directa o un indicio, la cuestión radica en la suficiencia de los mismos para satisfacer el estándar de prueba" (Vázquez, 2019, p. 206). Y, Jordi Ferrer Beltrán concluye en esta misma línea que "es tremendamente confundente llamar *presunción* a la conclusión de una inferencia probatoria, es decir, a un enunciado sobre los hechos que resulta probado por efecto de esa inferencia; oscurece más que aclara denominar *razonamiento presuntivo* al razonamiento probatorio; supone, seguramente, no haber percibido que todo el razonamiento empírico es inductivo, en cualquier área del conocimiento, de modo que si a ese tipo de inferencias las denominamos presunciones deberemos concluir que todo el razonamiento científico es presuntivo y buena parte también del que realizamos en nuestra vida ordinaria. Pero una categoría que sirve para tantas cosas temo que no sirva realmente para nada. (…) en efecto, si el razonamiento presuntivo no es nada distinto del razonamiento probatorio, decir que una de las maneras en que se puede probar un hecho es a través de las presunciones *hominis* o simples no es más que reducir el todo a una parte. No hay otra forma de probar el acaecimiento de un hecho del mundo que la de inferirlo a partir de otros hechos y de generalizaciones empíricas" (Ferrer Beltrán, 2022a, p. 424-425).

[53] También en el campo del derecho procesal civil, la distinción comúnmente imaginada es objeto de crítica intensa. Solo para ejemplificar, menciona Bruno Cavallone que "la considerazione del maggior rischio di errore insito nel ragionamento presuntivo (o meglio: nei più complesso tra due ragionamenti que tienen identica natura 'inferenziale') —considerazione senza dubbio esatta in astratto, ovvero in termini probabilistici o statistici— non giustifica affatto la conclusione che, in concreto, un giudice al quale si offrano una prova 'presuntiva' e una prova 'diretta' (di una medesima proposizione di fatto, o di due proposizioni incompatibili), debba necessariamente attribuire alla prima un'efficacia inferiore a quella della seconda; poiché evidentemente l'apprezzamento dell'una e dell'altra non è condizionato soltanto dalla loro struttura logica, ma anche e soprattutto, nel primo caso, dalla validità della 'regola di esperienza' che consente il passaggio dal fatto noto al fatto ignoto, e, in entrambi i casi, dall'attendibilità della fonte (il 'fatto percepito dal giudice') dalla quale si diparte, breve o lunga che sia, la catena 'inferenziale'" (Cavallone, 1991, p. 372-373). En la misma línea, entre tantos otros, v., Schmitz, (2020, p. 283-285).

ción judicial. Por otro lado, tal como ocurre con cualquier otro medio de prueba, la viabilidad del empleo de esta fuente de prueba (o, si se prefiere, siguiendo la jurisprudencia brasileña, de este indicio) —y ahora volviendo al tema de la conducta de las partes como elemento de prueba— dependerá siempre del cumplimiento de los requisitos necesarios para justificar, de forma racional, la formación de la convicción a partir de su empleo. Así, es posible decir que, como todo indicio y como todo elemento de prueba en general, la prueba por la conducta de las partes se somete a los criterios generales de aceptación de las justificaciones judiciales en materia de prueba. Empleando conceptos comúnmente aplicados a la prueba indiciaria, se puede generalizar esta idea afirmando que es necesario que la justificación ofrecida deba satisfacer los llamados aspectos intensivo y extensivo.[54] Por el segundo criterio (extensivo), es imprescindible que la justificación dada a partir del empleo de la conducta de la parte como prueba sea apta para demostrar la *totalidad* del hecho a probar. Y, para satisfacer el otro criterio (intensivo), es necesario que se demuestre que de aquel hecho (el comportamiento del sujeto) no puede derivar otro hecho que no sea aquel que se desea probar. Si no hay *suficiente* indicación de estos elementos, cumplidos los criterios de estándares probatorios válidos para cada causa, la justificación ofrecida puede verse debilitada,[55] en la exacta proporción en que aumenta la distancia con aquellos dos criterios.

Por ello, como suele ocurrir con cualquier medio de prueba, en la mayoría de los casos, la conducta de las partes cobra fuerza cuando se suma a otros elementos de prueba y pierde su fuerza cuando se considera de forma aislada. Aun así, como se verá, podrá haber situaciones en las que otras circunstancias, ligadas al caso o a las necesidades de tutela, justifiquen que dicho comportamiento, considerado de forma aislada, sea un argumento válido para la formación del convencimiento judicial.

Desde otro ángulo, esta afirmación se vincula con otra cuestión, también esencial para el tema en examen. Es común hacer alusión a una clasificación entre pruebas directas e indirectas en otro contexto que no es exactamente el que distingue las pruebas sobre el hecho y las pruebas sobre un hecho externo, del cual se deduce el hecho (la diferencia entre indicio y prueba directa). En este otro contexto, las pruebas también se clasifican en directas e indirectas, pero para enfatizar la relación que se da entre el sujeto cognoscente y el hecho a probar. En las pruebas directas (como la documental y la inspección judicial), este contacto sería *inmediato*, mientras que, en las pruebas indirectas, sería *mediado* por otro sujeto o por otro hecho (tal como sucede en la prueba testimonial o pericial).[56] Esta clasificación —que ya fue muy útil en el pasado—

---

[54] La denominación es oriunda del derecho americano (v. James, Hazard Jr. y Leubsdorf, 1992, p. 349).

[55] Ob. loc. cits.

[56] A ese respecto, pondera Carnelutti que "la differenza fra i due tipi di prova sta nella *coincidenza* o nella *divergenza del fatto da provare* (oggetto della prova) e del *fatto percepito dal giudice* (oggetto della percezione); appunto la prova indiretta presenta il distacco dell'*oggetto della prova* dallo *oggetto della percezione*: il fatto sotto posto alla percezione del giudice non serve se non di mezzo alla conoscenza di lui" (Carnelutti, 1947, p. 64-65).

supone que las pruebas indirectas están más sujetas a ofrecer diferentes *versiones* del mismo hecho, en la medida en que hay un tercero que también participa en la relación de conocimiento, mientras que las pruebas directas aportarían mayor seguridad.[57] El análisis del comportamiento de las partes quizás podría calificarlo como prueba *directa* o *indirecta*, dependiendo de las circunstancias, lo que podría sugerir que, en ciertos casos, aportaría mayor o menor seguridad al juzgador.

Esta clasificación, sin embargo, esconde un grave defecto: toda prueba pasa por un juicio crítico para ser considerada. Y, así, incluso pruebas que podrían calificarse como «directas» en este contexto pasan por un análisis subjetivo y pueden ser percibidas de forma diferente, según el sujeto que las examina.[58] En otros términos, en la correcta percepción de Vitor de Paula Ramos, (2023, p. 76-77) las pruebas materiales —y la conducta de las partes puede ser así calificada— son *signos* que necesitan ser interpretados para adquirir significado.[59] Y este proceso de interpretación, que dependerá de diversos factores, añade no solo sentido, sino también valores a lo que fue objeto del análisis.

Consciente de esta circunstancia, es fácil ver que, a veces, la conclusión que se extrae del valor probatorio del comportamiento de las partes se funda en preconceptos y juicios de valor subjetivos, que difícilmente pasarían el tamiz de una justificación racional. Al fin y al cabo, ¿sería posible decir que, solo porque alguien es un pésimo conductor, habiendo sido responsable de diversos accidentes de automóvil, también es responsable de un accidente específico en el que se vio involucrado? ¿O que la relación de cierta empresa con conductas ilícitas sería indicativa de que violó de mala fe el contrato objeto de un determinado proceso?[60]

---

[57] Una vez más, según Carnelutti, "la eccellenza della prova diretta sulla indiretta non ha bisogno di essere posta in rilievo: quanto più vicino è il fatto da provare ai sensi del giudice tanto più la prova è sicura" (Carmelutti, 1947, p. 65).

[58] Observa al respecto Giovanni Tuzet que "se a prova representativa equivale à prova direta, pode-se responder que o problema está mal colocado, na medida em que mesmo a prova representativa ou direta exige um passo inferencial. Como diz Ubertis, a peculiaridade da prova representativa reside no caráter 'automático' da inferência entre o elemento e o resultado da prova. Com maior razão para nossa discussão se, como discutido em outro lugar, todo processo perceptivo envolve um processo inferencial mais ou menos 'automático': mesmo uma prova 'direta', como a inspeção de um lugar, envolvem inferências por parte do sujeito que percebe (categorização do que é percebido, seleção mais ou menos 'automática' dos aspectos relevantes, sua comparação com outras cognições, sua valoração etc.)" (Tuzet, 2024, p. 47).

[59] "Por tanto, ya es importante señalar algo: al contrario de lo que presupone el procesalismo y la doctrina en general, el proceso de *interpretación* de una imagen o un dibujo, así como de un texto, no es algo que pueda prescindir del intérprete: *cuando se imagina que el intérprete solo está viendo, utilizando una fotografía como equivalente de la percepción, se está, en realidad, permitiendo que el intérprete presente su interpretación como si fuera algo objetivo*. (…) Por tanto, el proceso de reconstrucción de sentidos involucra una serie de complejidades, de modo que, para trabajar con documentos, es esencial entender la relación que cada tipo de signo puede tener con la realidad, a fin de comprender, también, cómo se pueden utilizar como prueba" (Ramos, 2023, p. 76-77).

[60] O, aún, de forma más amplia, ¿sería posible sostener, actualmente, la validez de reglas como aquella insertada en el art. 405, § 3º, incisos I y II, del CPC brasileño de 1973, que decía ser sospecho-

Considerar este tipo de comportamiento pasado —de forma aislada— como elemento importante para evaluar la conducta presente de alguien o para extraer elementos de prueba de este o aquel hecho puede ser algo extremadamente peligroso. Al fin y al cabo, esto puede simplemente revelar sesgos o prejuicios que contaminan el razonamiento judicial que vincula el indicio al hecho a probar y pueden dar al comportamiento de las partes una dimensión que no debería tener.[61]

Con estas consideraciones, asume particular relevancia la justificación dada por el magistrado para el empleo de este elemento de prueba y, sobre todo, para motivar el razonamiento utilizado para vincular el hecho instructivo (indicio) al hecho a probar. Al tratarse de prueba atípica, no hay un razonamiento o valoración previa del legislador que guíe la conducta del juez. Por ello, es esencial que el juez explique de manera adecuada cuáles son las circunstancias de la conducta de la parte que formaron su convicción, cuál es el razonamiento por él empleado para deducir la existencia de argumentos de prueba y de qué modo todo esto es suficiente para autorizar cierta conclusión.[62]

Y, si esta fundamentación es esencial, también lo es garantizar la *participación* de las partes en el diálogo que debe preceder a la decisión. Consciente de que el análisis de la prueba —incluida la conducta de las partes— es algo que exige *interpretación*, lo que demanda la elección de una entre varias alternativas posibles, no se puede admitir en el ámbito del proceso que esta decisión ocurra sin la necesaria participación de las partes, asegurada por la garantía fundamental del contradictorio (arts. 9º y 10, del CPC brasileño). El diálogo con los sujetos interesados puede contribuir a percibir posibles sesgos, a evaluar las diversas versiones (a las que puede conducir la conducta procesal de la parte analizada) y a entender la perspectiva completa de la situación fáctica a demostrar. Esta participación, por lo tanto, no puede ser descuidada, bajo

---

sos para declarar como testigo "el condenado por crimen de falso testimonio, habiendo transitado en juzgado la sentencia" o "el que, por sus costumbres, no fuera digno de fe"?

[61] Como observa Schauer, aunque estas conductas pretéritas puedan indicar predisposiciones (*dispositions*) para la adopción de comportamientos futuros, la conducta humana es el resultado de la interacción entre esas predisposiciones y las situaciones específicas en que las personas son puestas (ob. cit., p. 214). Por eso, razonar exclusivamente con base en las predisposiciones puede llevar a conclusiones equivocadas, con base en razonamientos tendenciales (*propensity reasoning*), la mayoría de las veces vedados (Schauer, 2022. p. 212).

[62] En suma, como observa Taruffo, "em verdade, as provas circunstanciais não se ajustam facilmente a regras analíticas e a restrições formais, motivo pelo qual as garantias necessárias de equidade e de credibilidade devem fundar-se, com maior razão, na racionalidade das inferências extraídas pelo julgador. Dessa maneira, quando confiável, as provas circunstanciais poderão ter o mesmo valor probatório que qualquer outro tipo de prova" (Taruffo, 2014, p. 105). Por otro lado, como pondera Ferrer Beltrán, "se trata, pues, de que la decisión probatoria cuente con buenas razones epistémicas y normativas que le den fundamento suficiente. Las primeras resultarán de la valoración individual y conjunta de la prueba, a los efectos de determinar el grado de corroboración que los elementos de juicio aportados al proceso otorgan a las distintas hipótesis fácticas en conflicto. Las razones normativas, en cambio, apuntan a la suficiencia o insuficiencia de esa corroboración, que habrá que justificar sobre la base de los estándares de prueba aplicables al caso" (Ferrer Beltrán, 2025b, p. 193).

pena de aislar al juez en una burbuja mental y de transformar el proceso de determinación de los hechos en una operación pobre, introvertida y miope, ajena a la riqueza que solo el contraste de visiones puede ofrecer.

Dadas estas condiciones, es importante notar que algunas condicionantes pueden contribuir a crear el ambiente adecuado para que la conducta de las partes pueda ser elevada a elemento de prueba suficiente para la determinación de los hechos relevantes del proceso.

6. MÓDULOS DE PRUEBA, CATEGORÍAS DE HECHOS Y LA PRUEBA POR LA CONDUCTA DE LOS SUJETOS

En un conocido estudio sobre la teoría de la prueba, Gerhard Walter (1985, p.167 e ss.) demuestra que las circunstancias del caso pueden ser altamente relevantes para determinar las posibilidades de prueba. Según el autor, no todas las situaciones permiten obtener el mismo grado de convicción.[63] Circunstancias específicas de la realidad impondrán límites a esta determinación de la realidad, de modo que sería irrazonable exigir un grado de convencimiento mayor que el que los límites concretos pueden ofrecer. Particularidades de las situaciones concretas —como la determinación de condiciones subjetivas, a ejemplo de objeciones de conciencia o de consecuencias que pueden derivar de varias causas distintas, como la epilepsia que puede haber sido desencadenada por un daño cerebral ocurrido muchos años antes o por alguna otra causa, quizás incluso desconocidad— son prácticamente imposibles de demostrar con el mismo grado de certeza con que se consigue demostrar elementos objetivos como la propiedad de un bien o la existencia de un contrato. En virtud de ello, el grado de convicción que se puede exigir debe respetar los límites de posibilidad de cada caso, bajo pena de hacer imposible la tutela de derechos en ciertas situaciones puntuales.

Es decir que *estándares* únicos o fijos de prueba — o, eventualmente, la adopción de solo dos modelos, uno para el proceso civil y otro para el penal — pueden ser insuficientes para explicar la riqueza de situaciones que ocurren en el mundo sensible.

Para ejemplificar la necesidad de *estándares* distintos para situaciones de cognición limitada, considérese el siguiente caso, juzgado en 2022, por la 1ª Sala del Tribunal

---

63 "Pero también en la constatación de elementos de figura legal 'normales', el módulo de la prueba se redujo en casos de dificultad de prueba. Recuérdense las indemnizaciones pagadas a víctimas del régimen nacionalsocialista y a refugiados, pero también los casos de control de conciencia de los objetores que se niegan a prestar el servicio militar. Por lo pronto, no resolveremos si desemboca también en una reducción del módulo de prueba toda la jurisprudencia relativa a la prueba por presunción simple. En suma, esta ojeada retrospectiva obliga a que nos preguntemos si en vista de la diversidad de los casos que se han de resolver y de los hechos que con tal fin han de constatarse, pero sobre todo por las consecuencias que esas decisiones acarrean, tiene sentido insistir en que el módulo de prueba sea unitario" (ob. cit., p. 172).

Regional Federal de la 3ª Región.[64] Un determinado apostador interpuso una acción contra la *Caixa Econômica Federal* (un banco público brasileño), alegando que había apostado en un concurso de LotoFácil y que su boleto había sido premiado. Decía que había hecho dos apuestas y que había cotejado el primero de los boletos con otro concurso de la misma lotería y que, al no haber coincidencia con los números, había tirado su boleto. Sin embargo, al cotejar el segundo boleto —que poseía los mismos números— verificó que los números de la apuesta fueron los sorteados, de modo que sería dos veces ganador del concurso. Sostenía que las circunstancias evidenciaban que él era el titular de este otro boleto premiado, pero que, por haber descartado el primer boleto, solo había recibido el premio correspondiente a un boleto. Juzgada improcedente la demanda, el Tribunal de apelación entendió por reformar la sentencia, concluyendo básicamente que los elementos indiciarios existentes deberían ser suficientes para concluir que el autor de la demanda era, realmente, titular del segundo boleto premiado. Dijo el Tribunal que los documentos anexados a los autos, «que corroboran la narrativa de los hechos expuesta en la inicial, la propia existencia de una apuesta ganadora —con 15 aciertos, así como la de f. 115— en la misma lotería y terminal en que el autor realizó las apuestas del primer lote, sin rescate del premio, como confirmado por la CEF (f. 138) permite concluir, con seguridad, que se trata de la apuesta realizada por el autor, cuyo boleto fue descartado tras la verificación incorrecta de los resultados».

Sin duda, no hay ahí un juicio de certeza absoluta sobre los hechos.[65] Tampoco exigió el tribunal alguna especie de certeza absoluta para la formación de su convicción. Si lo hubiera hecho, sin duda, jamás acogería pretensiones como esta, en la medida en que es imposible formar un juicio de certeza o de verdad respecto a la titularidad de boletos premiados extraviados. Casos así revelan que circunstancias excepcionales exigen la reducción del módulo de la prueba, de forma que disminuyan las expectativas de «posibilidad probatoria» y permitan que, con base en juicios de apariencia (cuando ese sea el máximo que se pueda obtener), sean apreciados con la prueba posible de producirse en esas circunstancias.

Más recientemente, y de forma más elaborada, parece que Jordi Ferrer Beltrán (2025a, p.277 e ss) llega a una conclusión similar. Analizando la compleja «prueba de la causalidad», concluye que el *estándar* de probabilidad preponderante es insuficiente para explicar la realidad del proceso, sugiriendo entonces la adopción de dos postulados para la determinación racional de la prueba de la causalidad. En primer lugar, «que la hipótesis sea la mejor explicación disponible de los hechos cuya ocu-

[64] TRF3, 1ª Turma. AC 0023493-29.2014.403.6100. Rel. Desembargador Federal Wilson Zauhy. DJe 23.08.22.

[65] Al respecto, observa Daniel González Lagier que "parece claro que la certeza absoluta no puede exigirse como requisito de la decisión judicial y que la inferencia probatoria no es válida o inválida como una cuestión de todo o nada, sino más o menos sólida. Cuanto más sólida, mejor confirmada estará la hipótesis. La solidez de la inferencia —y la confirmación de la hipótesis— es una propiedad graduable. (Gonzáles Lagier, 2022, p. 381).

rrencia se trata de probar, a la luz de los elementos de juicio existentes en el expediente judicial». Es decir, dependiendo de la estructura del derecho procesal aplicado, cabe evaluar si, entre las varias hipótesis presentadas por las partes (o, eventualmente, cuya especificación pueda corresponder al propio magistrado), aquella alineada con las pruebas disponibles en el proceso ofrezca la mejor explicación posible para los hechos relevantes para la solución de la controversia. En segundo lugar, debe evaluarse «el peso probatorio del conjunto de elementos de juicio relevantes incorporados al proceso sea tendencialmente completo (excluidas las pruebas redundantes)». En otras palabras, aquí interesa evaluar la completitud del material probatorio, favorable o desfavorable a cada parte, existente en el proceso y que permite la ponderación de estas varias hipótesis que pueden formarse (Ferrer 2025a. p. 2921-296).

Luego, el grado de exigencia en cuanto a la prueba y, en consecuencia, la capacidad virtual de la prueba de generar grados de convicción depende directamente de lo que las circunstancias concretas pueden ofrecer, bajo pena de crear exigencias probatorias absolutamente inalcanzables.

Esta percepción puede trasladarse al debate sobre el uso de la conducta de las partes como prueba en el proceso. Si, en algunos casos, la evaluación de esta conducta puede ser irrelevante y prescindible porque las circunstancias concretas permiten la obtención de pruebas más adecuadas y contundentes sobre los hechos, en otros la situación será diferente. Hay, sin duda, un universo de situaciones en las que pruebas directas, sólidas e indiscutibles sobre los hechos son inviables. Y para estas situaciones, la conducta de los sujetos pasa a tener un papel relevante.

En el ámbito de las *demandas preventivas* (inhibitorias),[66] por ejemplo, la formación de la convicción judicial por medio de la evaluación de la conducta pretérita de la parte es una medida necesaria, especialmente ante la imposibilidad de conseguir una prueba mejor para la decisión de la cuestión.[67] Aquí, la finalidad de la prueba es la determinación de hechos[68] que permitan «imaginar» lo que puede ocurrir en el *futuro*, algo absolutamente extraño al objeto tradicional de la prueba.

Supóngase la situación de una empresa que, ante cierta actividad anteriormente desarrollada, como el transporte de determinada carga peligrosa, se vio involucrada en cierto tipo de accidente por no cumplir con ciertas medidas de seguridad.

---

[66] En Brasil, se entiende por tutela inhibitoria la tutela *definitiva* que tiene por función impedir la violación *futura* de un derecho. Ella no depende de la prueba del daño o de la culpa, bastando la probabilidad del ilícito (art. 497, párrafo único, del CPC brasileño). No es una tutela cautelar o provisoria, sino una forma de tutela definitiva, pero preventiva.

[67] V., sobre eso, Arenhart (2003, p. 273 e ss).

[68] Lógicamente, el objeto de la prueba son siempre afirmaciones de hechos pretéritos. Sin embargo, *el objetivo final* de la prueba de esos hechos es permitir inferir que, en el futuro, cierta conducta puede ocurrir. La prueba, así, se puede decir, será de la "amenaza" de que en el futuro ocurra cierta lesión a un derecho. La conexión entre (la afirmación d) el hecho pretérito y la probabilidad de ocurrencia futura de la violación temida depende de un juicio que debe ser adecuadamente justificado por la decisión judicial.

Permaneciendo idénticas las circunstancias, para fines de tutela preventiva, parece razonable suponer que, si otra vez la empresa es admitida a prestar aquella actividad específica, su conducta pretérita indicará la presencia de riesgo de que un accidente similar ocurra nuevamente. La conclusión podría ser extraña e irracional, si se pensara para el ambiente tradicional de la tutela reparadora. Imaginar que el deber de reparar eventualmente impuesto ante el primer accidente debería imponer un nuevo deber de reparar un perjuicio aún no ocurrido solo por la persistencia de la actividad de riesgo puede sonar un completo absurdo. Sin embargo, para viabilizar la tutela preventiva del derecho e impedir la nueva ocurrencia de un acto ilícito, parece adecuado que el magistrado pueda adoptar medidas tendentes a *impedir la repetición del ilícito, o a forzar al demandado a tomar las precauciones que había descuidado, a fin de cohibir que ese ilícito pueda ocurrir*, lo que podría generar perjuicios y consecuencias graves a innumerables valores.

Considerando que, en el ámbito de la tutela preventiva, la cognición estará ligada a un juicio de «prognosis», orientado hacia el futuro, no cabe duda de que lo máximo que se conseguirá obtener son juicios de probabilidad o de apariencia. En este escenario, las restricciones sobre la prueba posible militan a favor del uso de la conducta pasada del sujeto como elemento importante para la determinación de hechos que puedan indicar la plausibilidad del interés a proteger o de la futura violación del mismo.[69] En efecto, exigir certeza como requisito para la concesión de la tutela inhibitoria es inviabilizar, en la práctica, esta forma de protección, que muchas veces, como en el caso brasileño, cuenta con recepción constitucional (art. 5º, inc. XXXV, de la Constitución). Al fin y al cabo, en esta esfera, es totalmente imposible para el magistrado prever con seguridad lo que ocurrirá en el futuro. En términos de tutela inhibitoria, por lo tanto, bastará que la parte actora demuestre que la *probabilidad de la ocurrencia de la violación de su derecho* —comparativamente con las evidencias existentes en sentido contrario— es mayor y esto será suficiente para ser admisible la protección frente a la amenaza de lesión. Y, siendo este el referente de cognición posible, la prueba por la conducta del sujeto asume un protagonismo singular.

Pero no solo para la demostración de *tendencias* futuras sirve la prueba por el comportamiento de los sujetos. Como otros mecanismos de prueba indiciaria, también puede servir para la evaluación de hechos presentes y pasados.

La prueba por el comportamiento antecedente de las partes (como, por cierto, admite la jurisprudencia y la doctrina tradicionales brasileñas en relación con la

[69] Tratando de las medidas cautelares, por cierto, también Jordi Ferrer Beltrán admite que se pueda trabajar con estándares probatorios inferiores al de la "probabilidad prevalente". Aunque limitando esa posibilidad a decisiones provisionales, de corta duración, o para las cuales la gravedad del error de no adoptar la medida pedida sea muy alta (en contraste con la gravedad del error de adoptarla), entiende el autor que, en ciertas circunstancias, es razonable —sobre todo a la luz de los valores en juego— emplear como *estándar* de prueba la mera circunstancia de una hipótesis "capaz de explicar los datos disponibles, integrándolos de forma coherente, y han sido confirmadas algunas de las predicciones de nuevos datos que la hipótesis permita formular" (Ferrer Beltrán, 2025b, p. 262).

prueba «indiciaria» en general) también puede adquirir importancia significativa en casos en los que la normalidad de las situaciones revela la intención de ocultar comportamientos que puedan tener algún grado de ilicitud o es de esperar que alguien haya actuado de mala fe o con finalidad abusiva o antijurídica. De cierta forma, aquí se puede recurrir a la lección de Parra Quijano (2000, p. 234), pensada para el contexto criminal, pero ampliamente aplicable también aquí. Como él pondera, «quien realiza un acto lícito voluntariamente y como medida de seguridad por imposición de la ley lo documenta, lo hace frente a testigos, o le es indiferente que se presencie su realización, aún en el mismo acto de simular, los documentos o comportamientos que se fingen son en cierta forma realizaciones voluntarias que permiten descubrir el acto oculto o la no realización del acto. Por el contrario, quien prepara la comisión de un delito procura hacerlo de tal manera que nadie lo presencie; sin embargo, por ser este un comportamiento humano que afecta en alguna forma la realidad deja huellas producidas en la comisión del mismo que permiten descubrirlo e identificar a su autor» (Parra Quijano 2000, p. 234).[70] Evidentemente, también el proceso civil puede tener que lidiar con comportamientos ilícitos e inmorales en innumerables circunstancias. Casos como las acciones de improbidad administrativa o incluso problemas de violencia doméstica son frecuentemente traídos a la consideración del proceso civil. Y en estos casos, como es claro, para la demostración de cuestiones que por regla se practican de forma oculta, no es razonable que el juez exija una prueba más contundente que la que las circunstancias concretas pueden ofrecer.

Por lo demás, de modo más general, estas conductas anteriores de los sujetos pueden ser muy útiles para la demostración de elementos subjetivos, necesarios para la incidencia de reglas jurídicas. La demostración de la buena o mala fe en general, de los motivos que llevaron a la celebración de cierto negocio jurídico, de vicios de la voluntad o de vicios sociales, por ejemplo, difícilmente puede obtenerse por medio de las pruebas tradicionales. Sin embargo, la evaluación de comportamientos antecedentes de los sujetos —o del propio comportamiento que precedió la celebración del negocio— puede contribuir, e incluso ser decisiva, para esta interpretación de los elementos que gravitan en torno a la relación jurídica objeto del proceso.[71]

---

[70] También en este sentido, y aún en materia penal, es la conclusión de Vito Gianturco respecto de la prueba indiciaria: "mentre nel campo del lecito e dei rapporti obbligatori vi sono prove precostituite (probationes probatae) o, almeno, prove dirette que possono agevolmente costruirsi nel processo, in materia penale, come argutamente rilevarono i pratici 'delicta non probantur per instrumenta' e, il più delle volte, manca la prova diretta del fatto e delle circostanze que l'accompagnano: il valore precipuo, poi, che nel diritto penale ha l'elemento intenzionale o psicologico, impone sempre, salvo in quei pochissimi casi, in cui veramente può dirsi que res ipsa in se dolum habet, un'indagine accurata intorno alla volontà del colpevole" (Gianturco, 1958, p. 1).

[71] En efecto, las *Federal Rules of Evidence* norteamericanas, tras excluir la posibilidad de que actos antecedentes (en particular, la personalidad del sujeto) puedan ser empleados como prueba de que alguien actuó de acuerdo con cierta índole en caso específico, establece que esa prueba de la personalidad puede ser empleada para otros fines, tal como "proving motive, opportunity, intente, preparation, plan, knowledge, identity, absence of mistake, or lack of accident" (regla 404 (b)(2), de las *Federal Rules of Evidence* de los Estados Unidos de América).

Además, aunque se pueda pensar de modo distinto en el ámbito penal —y, en específico, para la punición de alguien por delito de violencia doméstica— para la adopción de medidas civiles, como el alejamiento del hogar conyugal u otras medidas restrictivas de convivencia, basta la verificación de un *riesgo* a la vida o a la integridad de la persona, que puede ser determinada por el comportamiento (único o reiterado) del aparente agresor.[72] Aun cuando esta medida pueda restringir el derecho a la vivienda o a la propiedad de este supuesto agresor, el hecho es que sería mucho más grave permitir que este sujeto siga viviendo junto con la mujer, exponiendo a riesgo serio la vida o la integridad física de ella. Por ello, incluso si no hay prueba absoluta de las agresiones —lo que es muy común en el caso de agresiones morales, patrimoniales y psicológicas (art. 7º, inc. II, IV y V, de la Ley 11.340/06)— la simple conducta pretérita o actual del hombre —o incluso la simple palabra de la víctima respecto a la existencia de este riesgo, apoyada por otros elementos presentes en el proceso[73]— puede servir como evidencia suficiente para el otorgamiento de la medida protectora.

Este punto específico demuestra otro dato relevante para la apreciación del peso a atribuir a la prueba por la conducta de los sujetos: cuanto más importante es el bien jurídico a tutelar, menor tiende a ser el grado de exigencia de prueba y, consecuentemente, más fácil y razonable es la aceptación del empleo de esta fuente de prueba específica.[74] Volviendo al ejemplo anterior, muy probablemente la conducta preté-

---

[72] En esa línea, afirma de modo expreso el art. 12-C de la Ley 11.340/06 (conocida como Ley Maria da Penha) que "verificada la existencia de riesgo actual o inminente a la vida o a la integridad física o psicológica de la mujer en situación de violencia doméstica y familiar, o de sus dependientes, el agresor será inmediatamente alejado del hogar, domicilio o lugar de convivencia con la ofendida...".

[73] Es lo que expresamente prevé el art. 18, § 4º, de la Ley 11.340/06, que afirma que "las medidas protectoras de urgencia serán concedidas en juicio de cognición sumaria a partir del testimonio de la ofendida ante la autoridad policial o de la presentación de sus alegaciones escritas y podrán ser denegadas en caso de evaluación por la autoridad de inexistencia de riesgo a la integridad física, psicológica, sexual, patrimonial o moral de la ofendida o de sus dependientes".

[74] Parece ser semejante la conclusión a la que llega Jordi Ferrer Beltrán (2025b, p. 261). El autor aún añade un ejemplo que ilustra bien el embate entre la necesidad de búsqueda de evidencia suficiente y la necesidad de tutelar valores fundamentales contra daños irreparables. Dice él: "Imagínese, por ejemplo, que un niño pequeño es llevado por sus padres al servicio de urgencias médicas con lesiones óseas y hematomas. Los padres cuentan que el niño se cayó por las escaleras, pero el médico que atiende al niño considera probable la hipótesis de que el niño fuera víctima de maltrato familiar. En esas condiciones, da parte a la fiscalía de menores, quien toma el asunto y pide al médico una evaluación, a la luz de las pruebas médicas, de la probabilidad de que las lesiones sean producto de una caída por las escaleras y de la probabilidad de que sean fruto de maltrato. El médico concluye que, con las pruebas disponibles, la hipótesis del maltrato es menos probable, pero que no es descartable en absoluto y tiene también apoyo epistémico en ellas. Supóngase también, por seguir el ejemplo, que el médico añadiera que si el niño quedara ingresado en el hospital por tres días podrían hacerse pruebas diagnósticas, físicas y psicológicas, que resultarían cruciales para determinar el origen de las heridas. En esas circunstancias, ¿sería racional adoptar como medida cautelar el ingreso forzoso del menor y, en consecuencia, impedir que los padres pudieran llevárselo a su casa, a pesar de que la hipótesis del maltrato tuviera una probabilidad menor que la de la caída por las escaleras? Si la respuesta, como creo, es que hay situaciones en que esa decisión sería racional (de acuerdo con los valores en juego), ello implica que la probabilidad prevaleciente no es siempre el mínimo racional para decidir sobre los hechos" (ob. loc. cits.).

rita del supuesto agresor sería considerada como prueba insuficiente para autorizar su condena por un delito de violencia doméstica, ya que el valor predominante ahí sería su libertad. Sin embargo, si el objetivo de la tutela jurisdiccional es la vida o la integridad de la mujer —y, por lo tanto, si se debe o no conceder una medida protectora como el alejamiento del supuesto agresor del hogar conyugal—, ahí el grado de convencimiento es diferente y la conducta antecedente de esa persona (y también de la mujer) pasa a tener un peso relativamente mayor.

Cabe subrayar que —al contrario de lo que pueda imaginarse— esto no está necesariamente ligado a la naturaleza «cautelar» de este alejamiento o de otras providencias semejantes. Aunque las situaciones *provisionales* puedan armonizarse más fácilmente con este bajo grado de probabilidad, este elemento no debe ser elevado a la condición imprescindible para su aplicación. No se impone justificar esta menor exigencia de prueba solo con base en la supuesta conexión de esta tutela con la apariencia del derecho (*fumus boni iuris*). En otras palabras, aunque este también pueda ser un elemento relevante, aquí el aspecto que predomina es la *relevancia* (objetivamente considerada) del bien jurídico a proteger. Por lo tanto, incluso en situaciones que no tengan ninguna conexión directa con las ideas de *provisionalidad* o de *preventividad* de daños o ilícitos (con lo que muchos conectan la propia noción de tutela cautelar), la relevancia del bien jurídico protegido puede determinar la reducción del grado de convicción judicial exigido para la decisión, permitiendo que la conducta de los sujetos sustente, incluso de modo exclusivo, la convicción judicial. Así, por ejemplo, ocurrió con el caso que inicia este texto —referente a la responsabilidad de las empresas de transporte de carga por los daños ocurridos en carreteras— y ocurre en muchos otros casos, como la autorización para la adquisición de armas de fuego —para la cual el interesado debe demostrar su «idoneidad»—, al no responder a una investigación policial o proceso criminal, su «ocupación lícita» y su «aptitud psicológica», entre otras cosas (art. 4º, de la Ley 10.826/03) —o con la prohibición de contratar con el Poder Público brasileño— con la exclusión, entre otras situaciones, de empresas que en los últimos cinco años, hayan sido condenadas judicialmente, con sentencia firme, «por explotación de trabajo infantil, por sometimiento de trabajadores a condiciones análogas a la esclavitud o por contratación de adolescentes en los casos prohibidos por la legislación laboral» (art. 14, inc. VI, de la Ley 14.133/21).

Tampoco se puede dejar de considerar que conductas pasadas —pero ahora no siempre de los propios sujetos involucrados en un conflicto determinado— pueden ser (y son) consideradas para el relleno de varios conceptos abiertos, a ejemplo de la noción de «buenas costumbres» (v.g., art. 187, del Código Civil brasileño), de «diligencia habitual» (art. 667, del Código Civil brasileño), o de «prácticas de crédito responsable» (art. 6º, inc. XI, del Código Brasileño de Defensa del Consumidor).

Determinar lo que significa esa «diligencia habitual», por ejemplo, implica examinar las conductas que normalmente se practican, a fin de establecer lo que podría ser normalmente exigible en ciertas circunstancias. A partir de ahí, determinado ese

estándar de comportamiento, es posible contrastarlo con el caso concreto y evaluar si ese límite fue o no observado. En casos así, lo que se tiene es que se toman las conductas anteriores —quizás practicadas por el propio sujeto, quizás vistas a partir de un conjunto de situaciones que se insertan en la noción en cuestión (en el caso, la diligencia habitual)[75]— para el relleno del concepto vago empleado por la ley. Una vez rellenado este concepto, se regresa a la conducta (ahora la conducta concreta y específica, discutida en el proceso) para analizarla y aplicar o no la norma. Sin duda, en estos casos, las conductas pasadas —y ahora, eventualmente, incluso las conductas practicadas por terceros— llenan de sentido el concepto indeterminado y permiten, al juez y a las partes, controlar de forma racional la incidencia o no de la norma. Se trata, en fin, de una relevante función probatoria desempeñada por la conducta pretérita concretamente verificada.[76]

Finalmente, parece haber otro espacio a considerar en la evaluación de la utilidad de la conducta de las partes para la formación de la convicción judicial. Al inicio de este texto, se ponderó sobre la utilidad de la separación entre hechos concretos y cate-

---

[75] Así, por ejemplo, decidió el Superior Tribunal de Justicia, en relación con la actividad mercantil, y valiéndose de *comportamientos de las propias partes y de otros sujetos pertenecientes al mismo segmento*, que "*tendo em vista os usos e costumes do segmento empresarial e as práticas adotadas, de longa data, pelas partes*, encontram-se presentes os requisitos para que o silêncio reiterado das varejistas —sobre a adoção dos boletos bancários e o repasse do respectivo custo— seja considerado manifestação de vontade apta a produção de efeitos jurídicos, vale dizer: seja atestada a existência de consenso em relação à forma de pagamento das 'mercadorias' e à cobrança de tarifa. Isso porque: (i) configurado o comportamento negativo das farmácias e drogarias, que, por mais de dez anos, pagaram os boletos bancários sem manifestar qualquer insurgência contra tal modalidade de adimplemento e a respectiva tarifa que lhes era cobrada; (ii) inexiste controvérsia sobre o fato de ser prática corriqueira do segmento empresarial a comercialização mediante boletos bancários com o escopo de otimizar a logística de distribuição de medicamentos, cuja relevância pública decorre da Constituição de 1988 e da Portaria 802/98 do Ministério da Saúde; (iii) também é incontroversa a habitualidade das negociações celebradas entre a atacadista e as varejistas com a utilização da citada forma de pagamento; (iv) cabia às compradoras —sociedades empresárias cuja vulnerabilidade não se reconheceu nos autos— apresentar resistência contra o modo de adimplemento ofertado durante os longos anos da relação contratual, merecendo destaque o fato de ter sido pleiteada na inicial a manutenção do pagamento via boleto bancário, havendo apenas objeção acerca do repasse da tarifa; e (v) revela-se evidente a convicção da atacadista sobre a adesão das varejistas quanto às cobranças efetuadas, na medida em que beneficiadas com a agilidade da forma de pagamento e a consequente pronta entrega dos produtos, entre outras facilidades que lhes eram ofertadas" (STJ, 4ª Turma. REsp 1.580446/RJ. Rel. Min. Luis Felipe Salomão. DJe 25.03.21 – el énfasis no está en el original).

[76] Importante aplicación de esta idea se aplica al llamado derecho indígena. En Brasil, el tratamiento jurídico dado a indígenas puede variar según el grado de su integración a la sociedad civil (art. 4º de la Ley 6.001/73). Así, indígenas que estén íntegramente integrados a la sociedad civil pueden recibir un tratamiento prácticamente idéntico a no-indígenas, para los varios fines del Derecho. Lógicamente, el análisis de ese grado de adaptación depende, en gran medida, de la evaluación del comportamiento del indígena, de modo que el análisis de las consecuencias jurídicas de sus actos necesita tomar en cuenta su comportamiento pretérito. Así, por ejemplo, ya se consideró que un indígena que "está perfeitamente adaptado à sociedade civil, tendo suficiente compreensão dos usos e costumes nacionais, possuindo fluência na língua portuguesa", deve ser considerado como integrado, podendo responder por eventual delito cometido (STJ, 5ª Turma. RHC 141.827/MS. Rel. Min. Ribeiro Dantas. DJe 16.04.21).

gorías de hechos para el estudio en examen. Como pondera Schauer, la prueba puede ser útil para la determinación de *un hecho específico o de una categoría de hechos*.[77] Y los mecanismos que pueden ser útiles para la demostración de categorías de hechos pueden no servir para la determinación de hechos específicos, a ejemplo de las estadísticas. Esto se debe a que, muchas veces, el razonamiento exigido para la formación de cierta convicción es un análisis probabilístico (*probabilistic assessments*), quizás incompatible con situaciones en las que se exija otro grado de evaluación.

Schauer (2022, p. 222-223) ilustra la idea con el ejemplo de un jugador de tenis profesional. Supongamos que ese jugador, en juegos anteriores, fingió cierta lesión. ¿Debe esta circunstancia recibir alguna consideración, por parte del árbitro o del médico que atiende el evento, en el análisis de una alegación presente de que se ha lesionado en un determinado juego? El autor presenta entonces otra hipótesis. Suponiendo el mismo tenista profesional, imagínese que el árbitro o el médico sabe que los tenistas en general suelen fingir lesiones durante los partidos. ¿Sería distinta esta nueva situación?

En síntesis, ¿el conocimiento de que es común que los tenistas (en general) finjan lesiones puede servir de prueba de que un determinado tenista, en las mismas circunstancias, finge cierta lesión? Esta es la cuestión propuesta por Schauer, quien entonces observa que, traspasados los límites de los hechos específicos, el valor probatorio dado a estas consideraciones generales variará según las proporciones de la clase en que cierta persona o cierto evento es colocado.[78]

Se entra así en el campo del valor que las conductas pasadas —no necesariamente del sujeto específicamente considerado— pueden componer datos estadísticos que

---

[77] "We often use evidence to reach a conclusion about a specific fact or specific act. Did Susan rob the First National Bank on September 30? Was Thomas Jefferson the biological father of Sally Hemings's son Eston Hemings? How many votes for president did Donald Trump and Joe Biden each receive in the state of Michigan in November 2020? But just as often we use evidence to support or challenge a general hypothesis about a category of acts or events, or about some larger phenomenon. What do we (or scientists) mean when we or they say that cigarette smoking causes cancer? And what is the evidence for that conclusion? Are Volvos reliable? How do we know? Does the use of aerosol cans damage the ozone layer? Does damage to the ozone layer cause climate change? Does the increased legal availability of guns increase the incidence of unlawful gun-produced harm? Does playing violent interactive video game causa an increase in the aggressive tendencies of teenage males who play those games? Does an increase in aggressive tendencies among teenage male cause teenage males to commit actual acts of violence? Does the amount of alcohol consumption that would ordinarily have no detrimental effects create problems when consumed by pregnant women?" (Schauer, 2022, p. 48-49).

[78] «Once we move beyond individuals and their past individual acts, the evidentiary weight to be given to evidence of this sort will plainly vary with the size of the class within which we place this particular person or event. If the particular player, the legitimacy of whose injury we are trying to evaluate, is a thirty-eight-years-old male from some particular country, then statistics indicating that tennis players with these exact characteristics and from that country have a high incidence of fabricated injuries would seem to have a great deal of evidentiary weight, whereas statistics indicating that men have a higher rate of injury fabrication than do all players, regardless of gender, would be less weighty» (*idem, ibidem*, p. 223).

permitan concluir por la mayor o menor probabilidad de que un comportamiento específico haya ocurrido o vaya a ocurrir. Y, en este punto, no parece haber duda de que estas conductas anteriores —a veces de la propia persona involucrada, a veces generales— pueden servir como elemento de prueba, en la exacta medida en que se debe aceptar el empleo de la prueba estadística en el proceso civil.[79]

Podría decirse lo mismo en el ámbito de la distinción, propuesta por Jordi Ferrer Beltrán, entre la causalidad general (probabilística) y la causalidad específica, para la determinación de la responsabilidad civil (Ferrer Beltrán, 2025a, p. 286-288). Como ejemplifica el autor, si es cierto que la afirmación de que fumar causa cáncer se demuestra de forma estadística, no hay cómo determinar, solo a partir de esas estadísticas, que el hecho de que alguien padezca cáncer tenga relación directa y necesaria con la circunstancia de haber fumado gran parte de su vida, de modo que una acción interpuesta contra la empresa tabacalera con esa base estaría condenada al fracaso. Sin embargo, como también señala el autor, sería diferente si el Estado presentara una demanda contra la empresa tabacalera, solicitando que fuera responsabilizada por el aumento de los gastos en salud (ante el aumento de casos de cáncer, causado por el tabaquismo).[80]

Por lo tanto, dentro de esta dimensión, el empleo de conductas pasadas como elemento para la creación de una «categoría de hechos» que sea válidamente sostenible puede ser (y ha sido) un uso útil de este elemento de prueba. Aunque pueda ser insuficiente para la demostración de la ocurrencia de un hecho específico,[81] su empleo para sustentar razonamientos probabilísticos y «hechos abstractos» parece bastante razonable. El aumento del riesgo de cáncer por exposición al amianto;[82] la inexistencia de evidencias científicas de la eficacia de la fosfoetanolamina sintética (píldora del cáncer)[83] y la inutilidad del internamiento de personas con trastornos mentales en instituciones asilares (como los Hospitales de Custodia y Tratamiento Psiquiátrico), por la falta de evidencia de su utilidad y el riesgo de violación de derechos fundamentales,[84] son ejemplos de hechos abstractos que ya fueron empleados en diversas demandas y decisiones judiciales. Construidos estos «hechos abstractos» a partir de comportamientos específicos verificados y reiterados, vienen siendo aplicados —como hechos notorios— en el ámbito jurisdiccional.

---

[79] V., sobre eso, Arenhart, (2019, *passim)*.

[80] «En cambio, todo cambiaría si la acción contra la tabaquera fuera presentada por el Estado alegando el aumento del gasto sanitario causado por el aumento de casos de cáncer, causado a su vez por el humo del tabaco. Más allá de la necesidad de probar el efectivo aumento del gasto sanitario, en este caso, la hipótesis a probar sería de causalidad general probabilística, por lo que el recurso a las frecuencias estadísticas estaría totalmente justificado» (*idem, ibidem,* p. 288).

[81] V. el ejemplo citado en la nota 14, anteriormente.

[82] V.g., STF, Pleno. ADIns 3.356. Redactor del fallo Min. Gilmar Mendes. DJe 30.11.17.

[83] STF, Pleno. ADIn 5.501. Relator. Min. Marco Aurélio. DJe 26.10.20.

[84] Resolución del Consejo Nacional de Justicia n. 487/2023.

7. ¿HAY ALGUNA CONCLUSIÓN POSIBLE?

En la conclusión de este texto, quedan más dudas que certezas. Parece claro que la conducta de los sujetos puede ser empleada para la demostración de hechos en el proceso civil. Sin embargo, la gran cuestión que subsiste es: ¿en qué límite es racional aceptar este elemento de prueba y qué profundidad de justificación se debe exigir para ello?

No hay duda de que la fundamentación de la decisión sobre la prueba es esencial para controlar de modo objetivo la legitimidad de la convicción judicial formada a partir de ese comportamiento. Sin embargo, los sesgos y prejuicios también pueden comprometer la propia justificación de la decisión judicial sobre la prueba, dispersando el peso atribuido a esa conducta pasada en la formación de la convicción judicial, sobre todo en casos similares. En otras palabras, es posible que jueces con comprensiones diversas del mundo acaben atribuyendo un valor diferente a la misma conducta procesal de las partes, todas justificadas a partir de matices externos examinados. [85]

Por otro lado, las especies de pretensión y las posibilidades de prueba que las circunstancias ofrecen hacen más digeribles o menos razonables ciertas conclusiones expuestas por el Poder Judicial en este campo.

Así, en el ejemplo inicialmente ofrecido en este texto, respecto a la situación de las empresas de transporte de carga, parece mucho más fácil sostener que la conducta pasada sirva para fines de ofrecimiento de tutela inhibitoria que para la imposición de responsabilidad civil. Esto se debe a que, mientras la primera forma de protección dispensa la demostración de daño y culpa (art. 497, párrafo único, del CPC), la segunda exige esa demostración, que es mucho más compleja cuando se toma en consideración solo el indicio derivado de la conducta pasada de los sujetos. [86]

Por lo tanto, factores ligados a la pretensión material objeto de protección, a las posibilidades de prueba que las circunstancias concretas ofrecen, los riesgos de que las

---

[85] Como recuerdan Kahneman, Sibony e Sunstein, "onde há julgamento, há ruído, e mais do que você imagina. (...) Como no exemplo dos papiloscopistas, quando sabemos o que os outros pensam, o viés de confirmação nos induz a formar precocemente uma impressão geral e a ignorar informações contraditórias. Os títulos de dois filmes de Hitchcock sintetizam bem isso: um bom tomador de decisão deve objetivar sempre manter 'a sombra de uma dúvida', nunca ser 'o homem que sabia demais'" (Kahneman, Sibony y Sunstein, 2021, p. 247 e 250).

[86] Con razón, argumenta Jordi Ferrer Beltrán que "hemos visto en el capítulo nueve las razones por las que un diseño procesal que sea compatible con el debido proceso debe contener reglas que determinen el umbral de suficiencia probatoria. Esta es precisamente la función de los estándares de prueba. Pero el umbral no es necesariamente único: una misma hipótesis, con las mismas pruebas y el mismo grado de corroboración, puede estimarse probada si el estándar de prueba es X y no probada si es X + 1. Esto explica, por ejemplo, que las mismas pruebas puedan ser suficientes para considerar probado un hecho en un tipo de proceso e insuficientes para otro tipo de proceso; o suficientes para adoptar una decisión intermedia durante el proceso —una prisión preventiva, por ejemplo—, pero insuficientes para la decisión final en sentencia" (Ferrer Beltrán, 2022b, p. 472).

presunciones formadas revelen prejuicios y sesgos y tantos otros factores, deben merecer atención en la elaboración de la justificación externada en el uso de esta prueba en el proceso. Como tantos otros medios de prueba, esta prueba demanda una atención redoblada en la motivación presentada en la decisión respecto a la prueba, ya sea en relación con la racionalidad del juicio extraído a partir de la conducta del sujeto en sí considerada (justificación individual en relación a la prueba), ya sea en lo tocante a la asignación de esa prueba en el conjunto probatorio presente.

Si es esta justificación la que aporta racionalidad sobre la prueba empleada y sobre el valor que se le otorga, es importante también que, al ofrecer esas razones, el magistrado sea consciente de los límites y potencialidades también de este medio de prueba, de modo que no se excluya su empleo cuando sea posible, ni se use cuando esto suene inadecuado. De todas formas, no hay razón para excluir, *a priori*, el uso de la conducta procesal de los sujetos como prueba en el proceso civil. El empleo de este medio, por cierto, puede llevar más allá la frontera de la determinación de los hechos en casos complejos y, particularmente, para las acciones preventivas, abriendo nuevas posibilidades para la tutela jurisdiccional en espacios hasta entonces considerados inhóspitos. La prueba, al final, constituye el gran factor de pavimentación de la actividad jurisdiccional frente a lo desconocido. Y, así, quizás este nuevo elemento pueda ofrecer nueva riqueza y nuevos horizontes, aportando otras luces a la práctica judicial.

8. REFERENCIAS BIBLIOGRÁFICAS

Alberto, M. (2005). Valor probatorio de la conducta en juicio. Un aporte más para su consideración como indicio y otras cuestiones más. In J. W. Peyrano (Coord.), *Valoración judicial de la conducta procesal*. Rubinzal-Culzoni.

Amaral, P. O. (2021). *Provas – Atipicidade, liberdade e instrumentalidade* (3a ed.). RT.

Arenhart, S. C. (2003). *Perfis da tutela inibitória coletiva*. RT.

Arenhart, S. C. (2019). A prova estatística e sua utilidade em litígios complexos. *Revista Direito e Praxis*, *10*(1).

Arenhart, S. C. (2025). *Relendo o princípio da demanda*. RT.

Cambi, E., & Hoffmann, E. (2011, noviembre). Caráter probatório da conduta (processual) das partes. *Revista de Processo*, *201*.

Carnelutti, F. (1944). *Sistema de derecho procesal civil* (N. Alcalá-Zamora y Castillo & S. Sentís Melendo, Trad.; Vol. I). Uteha.

Carnelutti, F. (1947). *La prova civile* (2a ed.). Ateneo.

Cavallone, B. (1991). *Il giudice e la prova nel processo civile*. CEDAM.

Davis, K. C. (1942). An approach to problems of evidence in the administrative process. *Harvard Law Review*, *364*.

Davis, K. C. (1955, noviembre). Facts in lawmaking. *Columbia Law Review*, *55*(7).

Dotti, R. F. (2020). *Tutela da evidência*. RT.

Faigman, D. L. (2008). *Constitutional fictions*. Oxford University Press.

Ferreira, W. S. (2014). *Princípios fundamentais da prova cível*. RT.

Ferrer Beltrán, J. (2022). La decisión probatoria. In J. Ferrer Beltrán (Coord.), *Manual de razonamiento probatorio*. Suprema Corte de Justicia de la Nación.

Ferrer Beltrán, J. (2022). La motivación sobre los hechos. In J. Ferrer Beltrán (Coord.), *Manual de razonamiento probatorio*. Suprema Corte de Justicia de la Nación.
Ferrer Beltrán, J. (2025). *Motivación y racionalidad de la prueba* (2a reimpr.). Grijley.
Ferrer Beltrán, J. (2025). *Prueba sin convicción* (2a ed.). Marcial Pons.
Gianturco, V. (1958). *La prova indiziaria*. Giuffrè.
González Lagier, D. (2022). Inferencia probatoria y valoración conjunta de la prueba. In Ferrer Beltrán. J (Coord.), *Manual de razonamiento probatorio*. Suprema Corte de Justicia de la Nación.
Grasso, E. (1973). Dei poteri del giudice. In E. Allorio (Dir.), *Commentario del codice di procedura civile* (Vol. I, No. 2). UTET.
Heñin, F. A. (2009, abril). Valoración judicial de la conducta procesal. *Revista de Processo*, *170*.
James Jr., F., Hazard Jr., G. C., & Leubsdorf, J. (1992). *Civil procedure* (4a ed.). Little, Brown and Company.
Kahneman, D., Sibony, O., & Sunstein, C. R. (2021). *Ruído – Uma falha no julgamento humano* (C. de A. Leite, Trad.). Objetiva.
Lopes, B. V. C. (2006). *Tutela antecipatória sancionatória*. Malheiros.
Marinoni, L. G. (2011). *Abuso de defesa e parte incontroversa da demanda* (2a ed.). RT.
Mazini, P. G. (2020). *Tutela da evidência*. Almedina.
Mitidiero, D. (2022). *Antecipação da tutela* (5a ed.). RT.
Montesano, L., & Arieta, G. (1997). *Diritto processuale civile* (2a ed., Vol. II). Giappichelli.
Parra Quijano, J. (2000, julio-septiembre). Prueba indiciaria en el código de procedimiento penal brasilero y código de procedimiento penal italiano y su complementación con países europeos y americanos. *Revista de Processo*, 99.
Ramos, V. de P. (2023). *La prueba documental* (L. Criado Sánchez, Trad.). Marcial Pons.
Righi, I. O. (1981). Eficácia probatória do comportamento das partes. *Revista da Faculdade de Direito da UFPR*, *20*.
Schauer, F. (2022). *The proof – Uses of evidence in law, politics, and everything else*. Harvard University Press.
Schmitz, L. (2020). *Presunções judiciais*. Revista dos Tribunais.
Taruffo, M. (2014). *A prova* (J. G. Couto, Trad.). Marcial Pons.
Theodoro Jr., H. (2003). Comentários ao novo código civil. In S. F. Teixeira (Coord.), *Comentários ao novo código civil* (2a ed., Vol. III, Tomo II). Forense.
Tuzet, G. (2024). *As razões da prova* (Caio Badaró Massena, Trad.). Marcial Pons.
Vázquez, C. (2019). Técnica legislativa del feminicídio y sus problemas probatórios. *Doxa – Cuadernos de Filosofía del Derecho*, *42*.
Vázquez, C., & López, M. F. (2022). La conformación del conjunto de elementos de juicio: admisión de pruebas. In J. Ferrer Beltrán (Coord.), *Manual de razonamiento probatorio*. Suprema Corte de Justicia de la Nación.
Walter, G. (1985). *Libre apreciación de la prueba*. Temis.
Woolhandler, A. (1988). Rethinking the judicial reception of legislative facts. *Vanderbilt Law Review*, *41*.

Referencias legislativas y jurisprudenciales

Argentina. Ley nº 23.950, [título oficial de la ley].
Brasil. Ley nº 6.001, de 19 de diciembre de 1973, Estatuto do Índio. (Diário Oficial da União, 21/12/1973).
Brasil. Ley nº 10.826, de 22 de diciembre de 2003, Estatuto do Desarmamento. (Diário Oficial da União, 23/12/2003).
Brasil. Ley nº 11.340, de 7 de agosto de 2006, Lei Maria da Penha. (Diário Oficial da União, 08/08/2006).
Brasil. Ley nº 12.529, de 30 de noviembre de 2011, Lei de Defesa da Concorrência. (Diário Oficial da União, 01/12/2011).

Brasil. Ley nº 14.133, de 1 de abril de 2021, Lei de Licitações e Contratos Administrativos. (Diário Oficial da União, 01/04/2021).
Código Civil. (Brasil).
Código Civil. (Italia).
Código de Proceso Civil. (Alemania).
Código de Proceso Civil. (Brasil).
Código de Proceso Civil. (Chile).
Código de Proceso Civil. (Francia).
Código de Proceso Civil. (Italia).
Código de Proceso Civil. (Perú).
Código de Proceso Civil. (Portugal).
Código de Proceso Civil da Nación. (Argentina).
Código de Proceso Penal. (Brasil).
Código Penal. (Brasil).
Corte Interamericana de Derechos Humanos. (2020, 1 de septiembre). *Caso Fernández Prieto y Tumbeiro vs. Argentina*. Disponible en https://www.corteidh.or.cr/docs/casos/articulos/seriec_411_esp.pdf
España. Ley 1/2000, de 7 de enero, de Enjuiciamiento Civil. (BOE, núm. 7, 08/01/2000, p. 575728).
Resolución do Consejo Nacional de Justicia n. 487/2023. (Brasil).
Superior Tribunal de Justicia. (2009, 19 de octubre). *REsp 697.078/BA*. Rel. Min. Desembargador Convocado Honildo Amaral de Mello Castro.
Superior Tribunal de Justicia. (2011, 24 de febrero). *REsp 1.120.746/SC*. Rel. Min. Nancy Andrighi.
Superior Tribunal de Justicia. (2021, 16 de abril). *RHC 141.827/MS*. Rel. Min. Ribeiro Dantas.
Superior Tribunal de Justicia. (2021, 18 de juño). *AgRg no HC 659.689/DF*. Rel. Min. Sebastião Reis Junior.
Superior Tribunal de Justicia. (2021, 25 de marzo). *REsp 1.580446/RJ*. Rel. Min. Luis Felipe Salomão.
Superior Tribunal de Justicia. (2023, 29 de agosto). *AREsp 1.581.491/PR*. Rel. Min. Raul Araújo.
Superior Tribunal de Justicia. (2024, 4 de diciembre). *REsp 1.908.497/RN*. Rel. Min. Teodoro Silva Santos.
Superior Tribunal de Justicia. (2024, 4 de diciembre). *REsp 1.913.392/MG*. Rel. Min. Teodoro Silva Santos.
Superior Tribunal de Justicia. (2025, 8 de abril). *AgRg no HC 958.143/MG*. Rel. Min. Ribeiro Dantas.
Superior Tribunal de Justicia. (2025, 15 de abril). *AgRg no REsp 2.181.502/SC*. Rel. Min. Rogério Schietti Cruz.
Superior Tribunal de Justicia. (2025, 25 de abril). *AgRg no AgRg no REsp 2.141.306/PR*. Rel. Min. Joel Ilan Paciornik.
Superior Tribunal de Justicia. (2025, 26 de mayo). *AgRg no AREsp 2.791.847/SP*. Rel. Min. Joel Ilan Paciornik.
Supremo Tribunal Federal. (1999, 10 de septiembre). *HC 77.987*. Rel. Min. Marco Aurélio.
Supremo Tribunal Federal. (2002, 22 de febrero). *HC 81.305/GO*. Rel. Min. Ilmar Galvão.
Supremo Tribunal Federal. (2017, 30 de noviembre). *ADIns 3.356*. Redactor del fallo Min. Gilmar Mendes.
Supremo Tribunal Federal. (2020, 26 de octubre). *ADIn 5.501*. Rel. Min. Marco Aurélio.
Supremo Tribunal Federal. (2024, 4 de marzo). *HC 169.788*. Rel. Min. Alexandre de Moraes.
Tribunal Regional Federal da 1ª Región. (2013, 23 de agosto). *AI 0056520-92.2012.401.0000/DF*. Rel. Desembargador Federal Carlos Moreira Alves.
Tribunal Regional Federal da 3ª Región. (2022, 23 de agosto). *AC 0023493-29.2014.403.6100*. Rel. Desembargador Federal Wilson Zauhy.
Tribunal Regional Federal da 4ª Región. (2018, 8 de febrero). *AC 5001629-36.2015.404.7214*. Rel. Desembargador Federal Rogério Favreto.
Tribunal Regional Federal da 4ª Región. (data indisponible). *AC 5018457-33*. Rel. Jueza Federal Marga Barth Tessler.

Quaestio facti. Revista Internacional sobre Razonamiento Probatorio
Quaestio facti. International Journal on Evidential Legal Reasoning
Sección: Ensayos
2026 I 10 pp. 327-359
Madrid, 2026
DOI: 10.33115/udg_bib/qf.i10.23209
Marcial Pons Ediciones Jurídicas y Sociales
© Alejandra Olvera Dorantes
ISSN: 2604-6202
Recibido: 11/08/2025 | Aceptado: 31/10/2025 | Publicado *online*: 28/01/2026
Editado bajo licencia Reconocimiento 4.0 Internacional de Creative Commons

MOTIVOS PARA REPENSAR LA VIOLENCIA POLÍTICA CONTRA LAS MUJERES POR RAZONES DE GÉNERO EN LA LEGISLACIÓN MEXICANA: UN ENFOQUE DESDE EL RAZONAMIENTO PROBATORIO

Alejandra Olvera Dorantes

RESUMEN: La legislación mexicana prevé una infracción administrativa electoral que busca disuadir y sancionar la violencia contra las mujeres en el ámbito político. En este trabajo se exponen algunas de las complejidades probatorias que se presentan al intentar demostrar la comisión de esta conducta, y se abordan algunos de los principales postulados que ha sostenido la autoridad jurisdiccional para intentar resolver los problemas de prueba en el estudio de esta infracción. De modo específico se analiza la inversión de la carga de la prueba y el valor probatorio de la declaración de quien se aduce víctima, ambos criterios se contrastan con la manera en la que el Tribunal Electoral ha entendido los alcances de la perspectiva de género.

PALABRAS CLAVE: violencia política; razones de género; estereotipos; lenguaje; tribunal electoral; perspectiva de género; carga de la prueba.

REASONS TO RETHINK GENDER-BASED POLITICAL VIOLENCE AGAINST WOMEN IN MEXICAN LEGISLATION: AN APPROACH BASED ON EVIDENTIAL REASONING

ABSTRACT: Mexican legislation establishes an electoral administrative offense aimed at deterring and sanctioning political violence against women. This paper outlines some of the evidential challenges involved in proving the commission of such conduct and examines key doctrinal principles

adopted by judicial authorities in addressing the difficulties of proof associated with this offense. Specifically, it analyzes the reversal of the burden of proof and the evidential value of the statement made by the alleged victim. Both criteria are contrasted with the way in which the Electoral Tribunal has interpreted the scope and implications of the gender perspective.

KEYWORDS: political violence; gender-based grounds; stereotypes; language; Electoral Tribunal; gender perspective: burden of proof.

SUMARIO: 1. INTRODUCCIÓN Y PLANTEAMIENTO DEL PROBLEMA.— 2. CONCEPTO LEGISLATIVO Y JURISPRUDENCIAL DE LA VPMG: 2.1. Regulación de la VPMG; 2.2. Razones de género; 2.3. Actos u omisiones: *2.3.1. Cuando las conductas se dirijan a una mujer por su condición de mujer; 2.3.2. Afectación desproporcionada e impacto diferenciado*; 2.4. Lenguaje estereotipado.— 3. ANÁLISIS DE LOS CRITERIOS PROBATORIOS DEL TRIBUNAL ELECTORAL: 3.1. Juzgar con perspectiva de género como método de análisis; 3.2. Inversión de la carga de la prueba; 3.3. Consideraciones en torno a la declaración de quien denuncia VPMG.— 4. CONCLUSIONES.— V. FUENTES DE CONSULTA

1. INTRODUCCIÓN Y PLANTEAMIENTO DEL PROBLEMA

La violencia política contra las mujeres por razones de género —en adelante VPMG— está regulada en México como una infracción administrativa electoral[1] y surgió con la intención de contrarrestar el incremento de este tipo de violencia como consecuencia de la aplicación de acciones afirmativas para garantizar la participación paritaria de las mujeres en cargos públicos[2]. Actualmente es una de las conductas que más se analiza en las sentencias del Tribunal Electoral del Poder Judicial de la Federación —en adelante Tribunal Electoral o autoridad jurisdiccional—.

La VPMG se dirime a través del *procedimiento especial sancionador*[3] —en adelante PES—, que tiene una naturaleza administrativa sancionadora electoral y le rigen los principios del *ius puniendi* aplicables al derecho penal[4]. Sin embargo, en el estudio de esta conducta generalmente se presentan diversos problemas de relevancia, interpretación, calificación y prueba. Sobre los problemas de prueba, una de las dificultades

[1] Esta conducta también puede tener repercusiones en el ámbito penal, en términos del artículo 20 bis de la Ley General en Materia de Delitos Electorales; sin embargo, el estudio de este trabajo se delimitará a la VPMG en su vertiente de infracción administrativa electoral.

[2] Así lo establece, por ejemplo, la exposición de motivos de la Ley Modelo Interamericana para Prevenir, Sancionar y Erradicar la Violencia contra las mujeres en la vida política.

[3] Se precisa que el PES era resuelto en la primera instancia federal por la entonces Sala Regional Especializada del Tribunal Electoral; sin embargo, a partir de la reforma de 15 de septiembre de 2024 en materia de Poder judicial, dicho órgano jurisdiccional se extinguió y ahora es resuelto en única instancia por la Sala Superior del mismo Tribunal.

[4] De conformidad con la tesis XLV/2002 de rubro: «Derecho administrativo sancionador electoral. Le son aplicables los principios del *ius puniendi* desarrollados por el derecho penal».

es por la manera en que está regulada la VPMG en la legislación, y la forma en la que ha sido interpretada por la jurisprudencia del Tribunal Electoral. Esto, porque exige como elemento subjetivo para que se colme la infracción que la conducta esté basada en elementos de género.

Si alguien pregunta ¿cuándo se acredita que un acto u omisión está basado en elementos de género? La respuesta del Tribunal Electoral ha sido consistente: cuando se está en presencia de estereotipos de género. Dicha respuesta lleva a más interrogantes, ¿cuál es la línea que traza la distinción entre una acción, omisión u expresión que constituye un estereotipo de género y una que no lo es? O bien, cuando se trata de acciones u omisiones ¿cómo acreditar si omiten convocar a una Regidora a las sesiones de Cabildo de un Ayuntamiento por su género y no, por ejemplo, porque pertenece a un partido de oposición?

Otra de las dificultades, según ha sostenido en forma reiterada el Tribunal Electoral, es que las denuncias de esta infracción se dan en una situación de desequilibrio entre la denunciante —una mujer, dada la regulación de la conducta— y la persona denunciada —que puede ser cualquier persona—. Esto, al considerar que la VPMG involucra discriminación, por lo que se requiere de una tutela reforzada. Para atender esta cuestión, la autoridad jurisdiccional utiliza la perspectiva de género como método de análisis.

En ese sentido, el Tribunal Electoral ha establecido diversos criterios probatorios. Por ejemplo, ha dicho que es aplicable la figura de la *inversión de la carga de la prueba*. También reiteradamente ha dicho que «[…] los actos de violencia basada en el género pueden tener lugar en espacios privados donde ocasionalmente sólo se encuentran entre la víctima y su agresor y, por ende, no pueden someterse a un estándar de imposible prueba, por lo que su comprobación debe tener como base principal el dicho de la víctima leído en el contexto del resto de los hechos que se manifiestan en el caso concreto»[5].

Cabe precisar desde este momento que no es una situación menor determinar la comisión de VPMG, pues las sanciones que dan lugar a este tipo de conductas van desde amonestaciones públicas[6], sanciones pecuniarias, medidas de reparación del daño[7], la inscripción en un registro de personas sancionadas[8], dadas determinadas

[5] Véase sentencia: SUP-REC-200/2022, par. 68. Todas las sentencias que se citan en este trabajo son públicas y pueden ser consultadas en la siguiente liga electrónica: *https://www.te.gob.mx/buscador/*.

[6] Como ejemplo relevante véase la sentencia de la Sala Regional Especializada identificada con la clave SRE-PSC-94/2024, confirmada por la Sala Superior en la diversa SUP-REP-401/2024 y acumulados, las cuales tuvieron relevancia mediática con motivo de la sanción impuesta a una ciudadana consistente en disculparse públicamente con una diputada federal en la red social *X* durante 30 días consecutivos.

[7] Generalmente se le instruye a la persona infractora que ofrezca una disculpa pública, tome cursos de capacitación y difunda una síntesis de la sentencia.

[8] Registro Nacional de Personas Sancionadas, disponible para consulta en la siguiente liga electrónica: *https://www.ine.mx/actores-politicos/registro-nacional-de-personas-sancionadas/*.

circunstancias, la nulidad de elecciones[9] y, anteriormente, la posible inhabilitación para contender a cargos públicos[10].

Para analizar este escenario, la estructura del presente trabajo se compone de la siguiente manera. En primer lugar, se expondrá la regulación de la VPGM en la en la legislación y la forma en que se interpreta en la jurisprudencia, haciendo especial énfasis en la metodología que utiliza el Tribunal Electoral para resolver este tipo de asuntos. En segundo lugar, se analizarán las complejidades para acreditar *las razones de género.* Finalmente, con las herramientas que ofrece el razonamiento probatorio se analizarán los criterios probatorios que, como se ha dicho, ha establecido el Tribunal, esto con el objetivo de determinar si realmente cumplen o no con el cometido de resolver los problemas de prueba dentro de un sistema de garantías procesales.

2. CONCEPTO LEGISLATIVO Y JURISPRUDENCIAL DE LA VPMG

2.1. Regulación de la VPMG

De manera previa a que la VPMG se regulara en la legislación, diversas autoridades emitieron un «Protocolo para la Atención de la Violencia Política contra las Mujeres por Razones de Género» —en adelante Protocolo de VPMG—. En este se establecieron cinco elementos indispensables para determinar si se actualiza o no la conducta, a saber:

1. El acto u omisión se base en elementos de género, es decir: i. se dirija a una mujer por ser mujer; ii. Tenga un impacto diferenciado y desventajoso en las mujeres; y/o iii. Las afecte desproporcionadamente.

2. Tenga por objeto o resultado menoscabar o anular el reconocimiento, goce y/o ejercicio de los derechos político-electorales de las mujeres.

3. Se dé en el marco del ejercicio de derechos político-electorales o bien en el ejercicio de un cargo público.

4. Sea simbólico, verbal, patrimonial, económico, físico, sexual y/o psicológico.

5. Sea perpetrado por cualquier persona o grupo de personas.

Este documento, en principio, tenía el carácter de *soft law*; sin embargo, a raíz de diversos precedentes del Tribunal Electoral[11], el estudio de estos cinco elementos

[9] Esto solo para el caso de que se acredite que la VPMG fue determinante para el resultado del proceso electoral. Al respecto, véase la sentencia: SUP-REC-2214/2021.

[10] Se precisa que el 7 de marzo de 2023 la Suprema Corte de Justicia de la Nación resolvió en la contradicción de criterios 228/2022 que el concepto «modo honesto de vivir» no puede ser considerado un requisito de elegibilidad, pues determinó que un régimen constitucional democrático se debe rechazar la idea de un modelo único de moralidad.

[11] En específico, de las sentencias identificadas con las claves: SUP-JDC-383/2017, SUP-REP-252/2018 y SUP-REP-250/2018.

se convirtió en *hard law*, pues en la Jurisprudencia 21/2018, de rubro: «Violencia política de género. Elementos que la actualizan en el debate político», se determinó que quien juzga debe analizar si se colman o no estos cinco elementos. Así, desde antes de que la conducta estuviera regulada en la legislación, esta era punible a través de la jurisprudencia.

Posteriormente, con la reforma en la materia, se definió la VPMG en la Ley General de Acceso de las Mujeres a una Vida Libre de Violencia (2017) —en adelante Ley de Acceso—, como: «toda acción u omisión, incluida la tolerancia, basada en elementos de género y ejercida dentro de la esfera pública o privada, que tenga por objeto o resultado limitar, anular o menoscabar el ejercicio efectivo de los derechos políticos y electorales de una o varias mujeres, el acceso al pleno ejercicio de las atribuciones inherentes a su cargo, labor o actividad, el libre desarrollo de la función pública, la toma de decisiones, la libertad de organización, así como el acceso y ejercicio a las prerrogativas, tratándose de precandidaturas, candidaturas, funciones o cargos públicos del mismo tipo» (Artículo 20 bis de la Ley de Acceso).

La Ley dispone que se entenderá que las acciones u omisiones se basan en elementos de género cuando: a) se dirijan a una mujer por su condición de mujer; b) le afecten desproporcionadamente; o c) tengan un impacto diferenciado en ellas.

En el artículo siguiente (artículo 20 ter de la Ley de Acceso) se enuncia un listado de conductas que se subsumen en el supuesto de VPMG. Este catálogo se basa en la Ley Modelo Interamericana para Prevenir, Sancionar y Erradicar la Violencia contra las mujeres en la vida política —en adelante Ley Modelo— que, a su vez está inspirada en la normativa boliviana, pionera en legislar la VPMG[12]. Sin embargo, tal y como está establecido en la Ley de Acceso, se desprende que el listado no tiene pretensiones de taxatividad, sino que es únicamente enunciativo, puesto que la fracción XXII prevé que será constitutivo de la infracción cualesquiera otras formas análogas que lesionen o sean susceptibles de dañar la dignidad, integridad o libertad de las mujeres en el ejercicio de un cargo político, público, de poder o de decisión, que afecte sus derechos políticos electorales.

El artículo 442 bis de la Ley General de Instituciones y Procedimientos Electorales (2014) —en adelante Ley Electoral— también establece un listado enunciativo de conductas a través de las cuales puede manifestarse la VPMG, entre las que se encuentran: a) obstaculizar a las mujeres los derechos de asociación o afiliación política; b) ocultarles información para impedir la toma de decisiones; c) ocultar la convocatoria para el registro de candidaturas o información relacionada; d) proporcionar información falsa, incompleta o imprecisa para impedir su registro; y e) obstaculizar las campañas, impidiendo que se desarrollen en condiciones de igualdad.

[12] Ley Contra el Acoso y la Violencia Política Hacia las Mujeres (2012), disponible en: *https://www.ilo.org/dyn/natlex/docs/ELECTRONIC/90299/104007/F226460565/BOL90299.pdf*.

En la práctica cotidiana del Tribunal Electoral, generalmente se analiza la infracción realizando un *test*, en el que se estudia si se colman o no cada uno de los cinco elementos de la jurisprudencia referida; esto lo sigue haciendo así aún después de que entró en vigor la reforma legislativa. Adicionalmente, el Tribunal ha llegado a examinar si se está en presencia de alguno de los supuestos del catálogo enunciativo de conductas, pero siempre siguiendo esa metodología. Esto quiere decir que, a pesar de que la legislación establece ciertas conductas que se subsumen dentro del supuesto normativo, la autoridad jurisdiccional siempre analiza si se acredita o no que el acto u omisión se basa en elementos de género, dado que es uno de los ítems que conforman el *test*.

2.2. Razones de género

Antes de entrar al análisis de los criterios probatorios del Tribunal, resulta importante abundar en uno de los tópicos más importantes cuando se habla de VPMG, y posiblemente el más complejo, esto es, el elemento subjetivo consistente en que la conducta debe estar basada en elementos de género. El propósito de este apartado es enunciar las complejidades que engloba este requisito para que se colme la conducta denunciada. Para tal efecto, es importante distinguir las dos modalidades en que suele denunciarse la VPMG. La primera modalidad, cuando se trata de actos u omisiones; y, la segunda, cuando el asunto versa sobre expresiones o imágenes editadas, coloquialmente conocidas como «memes», lo que el Tribunal ha denominado *lenguaje estereotipado*. En ambos supuestos se presentan distintas dificultades para demostrar las *razones de género*, como a continuación se expondrá.

2.3. Actos u omisiones

El primer supuesto a estudiar como conductas constitutivas de VPMG es cuando se trata de actos u omisiones. Por ejemplo, para el caso de que una Regidora de un Ayuntamiento denuncie que en forma cotidiana se omite convocarla a las sesiones de Cabildo. En esta situación, la pregunta es: ¿cómo acreditar si omiten convocarla por su género y no, por ejemplo, porque pertenece a un partido político de oposición?

Una primera respuesta sería analizar únicamente los enunciados sobre los hechos externos y omitir el estudio de las razones de género para hacer un ejercicio subsuntivo del artículo 442 bis, inciso b) de la Ley Electoral el cual establece que constituye VPMG ocultar información para impedir la toma de decisiones; pues, con independencia de que esto se deba a cualquier otra circunstancia —como pertenecer a un partido de oposición— la legislación sustantiva vigente lo identifica como una conducta constitutiva de VPMG. Esto resolvería el problema del caso concreto. Sin embargo, puede darse el caso que se denuncien conductas que no sean susceptibles de subsumirse dentro en los catálogos —no taxativos— de la Ley Electoral y la Ley de Acceso.

Por otro lado, el órgano legislativo sí ha especificado las circunstancias a partir de las cuales se entenderá que existen *razones de género*; es decir, cuando: a) se dirijan a una mujer por su condición de mujer; b) le afecten desproporcionadamente; o c) tengan un impacto diferenciado en ellas.

Cabe precisar que, conforme a la técnica legislativa empleada, se advierte que bastaría con la configuración de cualquiera de las tres causales previstas para actualizar *las razones de género*, ya que el órgano legislador las enunció mediante una conjunción disyuntiva. De esta manera, la afectación proporcional y el impacto diferenciado constituirían supuestos independientes, susceptibles de configurarse aun cuando la diferencia de trato no esté expresamente motivada en el género de la persona. Sin embargo, conforme a la metodología de análisis del TEPJF se ha adoptado un criterio, conforme al cual, se exige la acreditación concurrente de todas las causales para tener por actualizadas las *razones de género*.

A continuación, expondré cada uno de estos supuestos:

2.3.1. Cuando las conductas se dirijan a una mujer por su condición de mujer

En primer lugar, es necesario señalar que sostener que un acto u omisión se dirige a una mujer por el hecho de serlo, implica afirmar que dicha conducta está motivada por el género de la persona perjudicada. Esto supone que la razón por la cual el Presidente Municipal omite convocar a una Regidora a las sesiones de Cabildo, o que una persona realice actos de presión para que una candidata renuncie a su aspiración política, estriba precisamente en su condición de mujer. En estos supuestos, el Tribunal ha puesto el énfasis en identificar estereotipos de género en las conductas —actos u omisiones— para demostrar que se dirigen a una mujer por su condición de mujer[13].

Por estereotipos, tomando como base la clasificación de Federico Arena[14] (2022, p. 182 y ss.), participan tanto los descriptivos como los normativos externos. Los primeros pueden identificarse cuando se busca dar un mensaje a partir del cual se afirme como verdadero que de que las mujeres, como grupo, carecen de aptitudes para participar en política o desempeñarse en asuntos públicos, cuya consecuencia es negar sus habilidades políticas. Esta idea puede ser un estereotipo descriptivo y, en consecuencia, constituir VPMG en perjuicio de la persona o grupo de personas concreta a quien se dirija este estereotipo, ya sea porque no existe un sustento empírico que justifique dicha premisa o porque en caso de estar sustentada en datos empíricos, estos son mal interpretados, o se muestra un error epistémico como reflejar mal la evidencia o estar integrada en una comprensión defectuosa del mundo[15].

---

[13] Por ejemplo, véase la sentencia SUP-REC-91/2020.

[14] Se precisa que el autor habla de estereotipos en términos generales, no particularmente estereotipos de género.

[15] Sobre evidencia mal interpretada o integrada en una compresión defectuosa del mundo, la filósofa Georgi Gardiner (2018, pp.18-22) expone dos ejemplos que me parecen ilustrativos. El primero,

Por su parte, los estereotipos normativos externos (Arena, 2022, pp. 182 y ss.) participan cuando se atribuye un rol o un deber a los miembros de un grupo, solo por pertenecer al grupo, y estos generan opresión, pues se pretende imponerle determinados atributos a ciertas personas o colectivos —en contraste con los estereotipos normativos internos, que generan reconocimiento porque las personas que sostienen el estereotipo a su vez son las destinatarias—. En el caso concreto cuando se realizan etiquetas como que las mujeres deben desempeñar el rol de amas de casa, mientras que a los hombres se les asigna el rol de líderes políticos[16].

En ese sentido, la infracción se configura si la persona agente considera que las mujeres carecen de aptitudes para desempeñarse en la política sobre la base de estereotipos de género y, a partir de ello, realiza actos u omisiones con la intención de obstaculizar la participación de la presunta víctima en asuntos públicos. Es decir, se trata de hechos internos que deben inferirse a partir de hechos externos (Gascón, 2005, p. 71; Vázquez, 2019, pp. 211-212)[17].

En este punto, es oportuno mencionar que, conforme a Daniel González Lagier (2022, p. 71), actuar intencionalmente es actuar con un propósito, persiguiendo un fin. Pero solo es posible decir que perseguimos un fin si sabemos dónde obtenerlo, o si confiamos en llegar a saberlo; es decir, si sabemos que hacer para lograr ese fin. A esta relación, el autor citado lo denomina *Principio de Racionalidad Mínima*, el cual se entiende de la siguiente manera «Si un agente actúa intencionalmente, tiende a realizar la acción que en ese momento cree más adecuada para lograr el fin que persigue».

Para ejemplificar lo anterior, se expone un caso real[18]. Se trata de una Síndica municipal que, además de ejercer su cargo público, se desempeñaba como maestra de prescolar por las tardes. La denunciante señaló que tras formular diversas solicitudes de información relacionadas con el manejo de recursos públicos, el Presidente Municipal realizó las siguientes conductas: emitió expresiones públicas denostativas tales como «Usted que está haciendo aquí, la están esperando los niños del Jardín»; ordenó la colocación de avisos visibles al público que instruían a la ciudadana acudir

---

cuando las personas sexistas consideran que el hecho de que las mujeres sean físicamente más débiles por término medio demuestra que las mujeres son inferiores a los hombres. El segundo, lo relaciona con la prueba estadística y refiere que: «la gente puede creer falsamente que la estadística indica que los negros tienen una mayor inclinación criminal por naturaleza, por ejemplo, en lugar de apreciar que la estadística indica que la marginación social y la opresión conducen a un aumento de los índices de delincuencia» (Gardiner, 2018, pp.18-22).

[16] Etiquetas que podrían ubicarse dentro de la clasificación de estereotipos que asignan roles, conforme a las autoras Cook y Cusack (2009, p. 29 y ss.).

[17] Las autoras coinciden en que esto es así siguiendo un modelo epistemológico cognoscitivista, entendiendo por aquel modelo el cual según los procedimientos de fijación de los hechos se dirigen a la formulación de enunciados fácticos que serán verdaderos si los hechos que se describen han sucedido y falsos en caso contrario (Gascón, 2005, p. 49). La profesora Carmen Vázquez, se hace cargo de exponer que existen textos normativos en realidad adscriben —es decir, no prueban— estados mentales, en ese punto su preocupación radica la posible infrainclusión y suprainclusión de los casos (Vázquez, 2019, pp. 211-217).

[18] Resuelto por la Sala Regional Monterrey del TEPJF en la sentencia: SM-JDC-389/2024.

directamente a la presidencia municipal para la atención de asuntos que correspondían a la esfera de actuación de la Síndica, a partir del siguiente contenido «Si busca a la síndico segundo por temas de panteón y no se encuentra puede pasar el Alcalde. Favor de registrarse primero en recepción. Gracias»; acudió de manera reiterada al centro educativo al que laboraba para solicitar su despido; y retiró de su oficina insumos indispensables para el ejercicio de sus funciones.

En ese sentido, si el Presidente Municipal tiene la idea generalizada de que las mujeres carecen de cualidades para desenvolverse en política y la suposición concreta de que la Síndica carecía de aptitudes para desempeñar su cargo a la par de actividades en el centro educativo, y a partir de ello realizó determinadas acciones u omisiones con el propósito de obstaculizar el ejercicio de sus funciones como funcionaria, dicha creencia opera como un estereotipo de género, a partir del cual, se estima incompatible las labores de cuidado y educación de infantes, con el desempeño de un cargo público.

Dicho estereotipo funciona como una premisa práctica en el razonamiento del agente, en el sentido de que le permite considerar como adecuados determinados medios para alcanzar el fin perseguido. Así, las conductas realizadas —hechos externos— permiten inferir racionalmente la intención de obstaculizar el ejercicio del cargo de la denunciante, en la medida que constituyen medios que el Presidente Municipal consideró idóneos a partir de la creencia estereotipada que orientó sus acciones.

Entonces, los estereotipos de género pueden cumplir dos funciones distintas. La primera, como una regla metodológica de no utilizarlos como instrumentos inferenciales, lo que la jurisprudencia ha establecido como obligación de cuestionar los hechos y valorar las pruebas desechando cualquier estereotipo o prejuicio de género, a fin de visualizar las situaciones de desventaja provocadas por condiciones de sexo o género [19].

El segundo, en un plano distinto, como la razón explicativa de la acción u omisión; es decir, como el elemento que explica la intención de la conducta, de tal manera que la razón por la cual el Presidente Municipal obstaculiza las funciones de la Síndica, puede estar sustentarse en la idea preconcebida que tiene respecto al rol de las mujeres en política, sin que ello implique confundir dicha explicación con la intención discriminatoria misma.

Por otra parte, la profesora Carmen Vázquez al hablar del tipo penal del feminicidio (2019, p. 214) precisa que, en su concepto, podría hablarse de dos posibilidades de incorporar los estereotipos de género al razonamiento probatorio: las máximas de la experiencia y el conocimiento experto.

Sobre el conocimiento experto, al menos en cuanto a aquel que proviene de una prueba pericial, quedaría descartado para los asuntos de VPMG —en el ámbito ad-

[19] Véase la jurisprudencia 1a./J. 22/2016 (10a.), de rubro: «Acceso a la justicia en condiciones de igualdad. Elementos para juzgar con perspectiva de género».

ministrativo electoral— ya que la Ley Electoral establece que en el PES únicamente será admitida la prueba documental y la técnica[20]. Es decir, no contempla pruebas personales. Esto tiene su fundamento en que el PES, en sus orígenes, se instauró como un procedimiento sumario, dispositivo y preventivo para evitar daños irreparables dentro de un proceso electoral.

En lo que respecta a las máximas de la experiencia, el profesor González Lagier (2005, p. 61) distingue entre aquellas de carácter científico o especializado, como las que aportan los peritos —se reitera que legislativamente la prueba pericial no es admisible en el PES—; las de carácter jurídico, como las derivadas del ejercicio profesional del juez o jueza; y las de carácter privado, esto es, provenientes de las experiencias de las y los jueces al margen del ejercicio de su profesión[21]. También, refiere que la estructura puede verse de la siguiente manera: Si «X» probable «Y». Así, estas generalizaciones fungirían como las «garantías» en el esquema de Toulmin, como lo mencionan tanto el profesor citado, como Anderson, Schum y Twining (2015, p. 96).

Un ejemplo que puede ilustrar la manera en que el Tribunal Electoral ha incorporado el uso de estereotipos de género a partir de máximas de la experiencia, y que se expone al margen de coincidir o no con la corrección de dichas inferencias, es la sentencia SUP-REC-91/2020 en el cual la Sala Superior confirmó la diversa SX-JDC-151/2020 de la Sala Regional Xalapa. En este asunto, una Regidora denunció, entre otras, no recibir sus prestaciones legales, no ser convocada a las sesiones, no recibir inmuebles para el desempeño de sus funciones, ser agredida verbalmente y no ser tomada en cuenta en las actividades del municipio, conductas oponibles al Presidente Municipal y diversos servidores públicos de Santa Lucía del Camino en Oaxaca.

Tanto la Regidora como el Presidente Municipal presentaron como prueba el mismo vídeo en el que se advierte una confrontación entre ambas partes; la presunta víctima pretendía acreditar que el funcionario la amenazó con denunciarla por fraude; mientras que el denunciado pretendía acreditar que la Regidora lo ofendió con insultos. Es decir, en este supuesto no eran los hechos brutos los que se cuestionaban, sino la percepción e interpretación de estos. En el vídeo se visualizaban a ambas partes discutiendo y que además había otro funcionario público en el lugar (párr. 205-208). Las conclusiones del órgano jurisdiccional en lo que respecta al análisis de los elementos de género, fueron las siguientes:

> Además de lo anterior, del video también se observa que el Presidente Municipal es acompañado de otro hombre, y entre los dos le hacen los señalamientos a la actora, lo que genera un actuar estereotipado de que los hombres unidos pueden invisibilizar o intimidar a una mujer, lo cual

[20] En términos del artículo 472, numeral 2 de la Ley Electoral. No pasa inadvertido que excepcionalmente el Tribunal ha admitido otro tipo de pruebas; sin embargo, ello no ha ocurrido así en forma reciente, pues así lo informó la autoridad instructora de los PES al responder una solicitud de información enviada para la presente investigación.

[21] En cuanto a aquellas de carácter jurídico y privado, en el contexto mexicano, yo añadiría que no solo se trata del conocimiento de quien juzga en sentido estricto, sino de las personas que colaboran en la elaboración de las sentencias.

> se evidencia ya que ambos hacen referencia al posible fraude y al inicio del procedimiento que quieren iniciar en su contra.
>
> Ahora bien, no pasa inadvertido para esta Sala Regional que al final del audio y del video se escucha y observa que la ciudadana […] le habla al ciudadano […] de forma irrespetuosa, con palabras altisonantes y reacciona de una manera agresiva; sin embargo, ello por sí mismo no desvirtúa el indicio de que recibe un trato discriminatorio por parte del Presidente Municipal del Ayuntamiento de Santa Lucía del Camino, Oaxaca y, mucho menos, que el Presidente Municipal sea la víctima.
>
> Lo anterior, porque como se observa, la actora reacciona de una manera explosiva, al ser cuestionada y señalada como posible responsable de un hecho presuntamente constitutivo de un delito y no obtener una respuesta a las peticiones que le ha formulado al Presidente Municipal y ante los señalamientos de que no la han convocado a eventos que supuestamente se han realizado en beneficio de mujeres, aun y cuando ella es la Regidora de Equidad, Género y Grupos Vulnerables.
>
> Además, como máxima de la experiencia, se estima que una persona que constantemente está sujeta a violencia ya sea física, psicológica, verbal o de alguna otra índole, en algún punto sí puede llegar a reaccionar contra su agresor; sin que este actuar le reste la calidad de agresor al perpetrador denunciado por la violencia política.

Así, para el Tribunal el que dos hombres estuviesen en el mismo lugar realizando señalamientos a la denunciante constituye el hecho externo a partir del cual se infiere un actuar estereotipado de su parte. En este razonamiento puede observarse, de forma entimemática, una máxima de la experiencia consistente en que, cuando dos hombres realizan cuestionamientos a una mujer, ello puede reflejar la idea de que, actuando conjuntamente, se refuerza una posición de autoridad o deslegitimación frente a ella, lo que se asocia con estereotipos de género sobre la subordinación de las mujeres en espacios públicos. Además, el Tribunal, sostuvo explícitamente que es una máxima de la experiencia que, cuando una persona se encuentra sujeta a una situación de violencia, puede reaccionar en contra de quien la agreda.

Con independencia de concluir o no con la justificación y la conclusión del Tribunal, la complejidad que se advierte en este tipo de asuntos radica en el control epistémico que debe ejercerse sobre las máximas de la experiencia que se invocan, a fin de evitar inferencias epistémicamente deficientes, ya sea al atribuir a los estereotipos un carácter explicativo de la acción u omisión, o al excluirlos como instrumentos inferenciales.

2.3.2. Afectación desproporcionada e impacto diferenciado

Es conveniente reiterar que la afectación desproporcionada y el impacto diferenciado son categorías que aluden a los efectos desiguales que una conducta puede producir sobre un grupo especifico —en este caso, sobre las mujeres—, las cuales, en principio, podrían configurarse aun cuando el acto u omisión no esté explícitamente motivado por el género de la presunta víctima. Sin embargo, el Tribunal Electoral ha establecido una metodología de análisis conforme a la cual, para tener por acreditadas las *razones de género*, se requiere la configuración de estos supuestos.

Sobre el impacto diferenciado, el Tribunal ha sostenido que este requisito se actualiza cuando la acción u omisión afecta a las mujeres *en forma diferente que a los hombres* o cuyas consecuencias se agravan por su condición de mujer. En cuanto a la afectación desproporcionada, para el órgano jurisdiccional, se trata de hechos *que afectan en mayor proporción a las mujeres que a los hombres* (sentencia SUP-REP-394/2021.).

Debe tomarse puntualizarse que a ambos supuestos les subyace una lógica común consistente en partir del principio de igualdad entendido en su dimensión material; es decir, superando la concepción formal de la igualdad —basada en una supuesta universalidad abstracta[22]— y reconociendo la necesidad de tutelar las desigualdades estructurales existentes para alcanzar una igualdad sustantiva o material[23].

En ese sentido, resulta pertinente introducir la noción de discriminación indirecta —o por resultado— la cual se presenta cuando una norma, práctica o criterio que aparenta ser neutral produce consecuencias adversas para determinado grupo social, es decir, cuando su aplicación genera un impacto diferenciado que da lugar a distinciones, restricciones o exclusiones no previstas de manera expresa (Serrano y Ortega, 2016, p. 39)[24].

En el caso específico de la VPMG, el análisis del impacto diferenciado y la afectación desproporcionada exige identificar si, a partir de una determinada acción u omisión, se genera una distinción injustificada en perjuicio de las mujeres y si, en consecuencia, la aplicación de criterio aparentemente neutral produce efectos desiguales y desproporcionadas sobre ellas, precisamente por las condiciones estructurales de desigualdad que enfrentan.

Ahora bien, en ocasiones, para acreditar estos elementos se acude a una argumentación generalizada, a partir de la cual se toma como premisas abstractas la desigualdad entre hombres y mujeres en el ámbito político; la situación general de violencia contra las mujeres en el país[25] o; cuando se presentan situaciones de interseccionalidad se alude a que dicha circunstancia las sitúa en un mayor grado de vulnerabi-

---

[22] Al respecto, véase: Barrère, M. Á. (2008). *Iusfeminismo y derecho antidiscriminatorio: Hacia la igualdad por la discriminación*. En R. M. Mestre i Mestre (Coord.), Mujeres, derechos y ciudadanías (pp. 45–72). Tirant lo Blanch.

[23] Sobre las diferentes maneras de entender a las desigualdades, véase: Ferrajoli, L. (2010). El principio de igualdad y la diferencia de género. En J. A. Cruz Parcero & R. Vázquez (Coords.), Debates constitucionales sobre derechos humanos de las mujeres (pp. 31–58). Suprema Corte de Justicia de la Nación; Fontamara.

[24] Un ejemplo de discriminación indirecta en materia político-electoral se observa en el diseño de acciones afirmativas, particularmente en aquellos sistemas electorales que reservan un porcentaje de candidaturas para mujeres electas por el principio de representación proporcional, pero que simultáneamente incorporan mecanismos como el denominado «mejor perdedor», mediante el cual las candidaturas derrotadas en el sistema de mayoría ocupan los primeros lugares de las listas de representación proporcional, lo que en la práctica termina desplazando a las mujeres y neutralizando los efectos de la acción afirmativa (Serrano y Ortega, 2016, p. 39).

[25] Véase la sentencia SRE-PSC-94/2024, confirmada por la Sala Superior en la diversa SUP-REP-401/2024.

lidad[26]. Asimismo, se toman en cuenta estadísticas como los índices de violencia contra las mujeres en el país[27], o hechos contextuales como: la cantidad de llamadas de emergencia que se recibieron relacionadas con incidentes violentos, la cantidad de reportes por agresiones a mujeres y la tasa de feminicidios en la entidad federativa que se trate.

Sin embargo, estos datos, por sí solos, no son ilustrativos para demostrar o contextualizar la violencia del caso concreto, puesto que no versan sobre VPMG y no están relacionados con la hipótesis de investigación[28].

Aunado a ello debe tomarse en cuenta una complejidad adicional, y es que en la VPGM es una conducta que se suscita en el ejercicio de los derechos políticos y electorales de las mujeres. Así, en el terreno de lo político y electoral, son naturales las circunstancias de confrontación, ya sea por la búsqueda del sufragio o dentro del ejercicio de los cargos de elección popular. Entonces, es innegable la desventaja histórica que han sufrido las mujeres que buscan acceder a cargos de representatividad, pues esto ha tenido que ocurrir mediante acciones afirmativas. Esta es la razón misma de establecer la VPMG como infracción.

Ahora bien, si para analizar la afectación desproporcionada y el impacto diferenciado se toma como punto de partida una concepción estructural del principio de igualdad[29], el énfasis no se coloca en la situación individual de la persona afectada, sino en las características específicas y los patrones comunes a partir de los cuales se ejercen determinadas conductas contra las mujeres como grupo históricamente en situación de desventaja, sometimiento u opresión. Desde esta óptica, es precisamente la identificación de dichos patrones, los cuales deben ser recurrentes en la arena político-electoral, lo que permite advertir si una acción u omisión ocurre por *razones de género* —en sentido amplio—. En ese sentido, no debe perderse de vista que estos patrones estructurales debieron ser identificados por el órgano legislativo al momento de regular la VPMG[30].

---

[26] Véase sentencia SRE-PSC-5/2023, párr. 198-203.

[27] Por citar algunos ejemplos, véase: SRE-PSC-17/2022 y SRE-PSC-118/2022.

[28] Sobre la manera en que se debe contextualizar un caso concreto, la herramienta proporcionada por la FLACSO para tales efectos (Ansolabehere, *et. al.*, 2017, pp. 34-35) indica que no todos los contextos son relevantes, de forma tal que, al responder la pregunta ¿cómo saber cuáles de esos hechos, conductas o discursos son relevantes? La respuesta es relativamente simple: ello está subordinado a la(s) hipótesis de investigación que se tengan.

[29] El principio de igualdad desde el punto de vista estructural es aquel que considera relevante la situación de la persona considerada, pero como integrante de un grupo sistemáticamente excluido. Esta visión, entiende que la igualdad ante la ley persigue el objetivo de evitar la constitución y establecimiento de grupos sometidos, excluidos o sojuzgados por otros grupos (Saba, 2007, pp. 167-187).

[30] Al analizar el tipo penal de feminicidio, particularmente cuando la técnica legislativa alude a las relaciones asimétricas de poder, la profesora Carmen Vázquez (2019, p. 210) hace hincapié en que la creación misma de la norma debió suponer un serio debate legislativo sobre si era necesario crear el tipo penal dadas ciertas características que la sociedad pretende regular, mismas que permitieran considerar las relaciones desiguales entre hombres y mujeres.

Así, aunque los catálogos enunciativos previstos en las leyes referidas no tienen pretensiones de taxatividad, cuando se denuncian conductas que pueden ubicarse dentro de dichos listados, bastaría, en principio, con realizar un ejercicio de subsunción normativa. No obstante, los problemas relacionados con la acreditación de las *razones de género* persisten cuando se denuncian conductas que no son susceptibles de subsumirse en esos catálogos, lo que plantea la duda sobre si resulta preferible adoptar un modelo estrictamente taxativo, como el previsto en la legislación boliviana o ecuatoriana, o bien el propuesto por la Ley Modelo. Esto, bajo la premisa de que el órgano legislativo ya habría asumido la tarea de identificar las conductas que reproducen patrones comunes de género en el ámbito político-electoral, lo que tendría como consecuencia que, en sede jurisdiccional, el análisis se limitara a los hechos externos denunciados, prescindiendo de un examen probatorio sobre la actualización de *las razones de género*.

2.4. Lenguaje estereotipado

El segundo supuesto de VPMG se presenta cuando se analizan expresiones o imágenes editadas —coloquialmente conocidas como «memes»— difundidas principalmente en redes sociales como *Facebook*, *X*, *Instagram*, *YouTube* o *TikTok*, así como en notas periodísticas o en propaganda político-electoral. Estas situaciones pueden ir desde un comentario en una red social, a campañas sistemáticas de desprestigio digital, las cuales han llegado incluso a provocar la renuncia de mujeres a sus candidaturas[31].

Hablar del lenguaje estereotipado en los casos de VPMG resulta relevante debido a las implicaciones que este tipo de expresiones tiene en relación con los *motivos de género*. Ahora bien, ¿qué ha considerado el Tribunal como lenguaje estereotipado? Lo cierto es que, cuando se denuncian expresiones o imágenes editadas, para que estas sean calificadas como VPMG deben reproducir estereotipos de género, conforme a la reiterada jurisprudencia del Tribunal Electoral[32].

Algunos casos en los que el órgano jurisdiccional ha establecido la existencia de leguaje estereotipado como constitutiva de VPMG puede ser la sentencia SRE-PSC-157/2021, en la cual se analizó la siguiente frase en *Facebook*: «Con todo respeto estas señoras ni como amas de casa, ninguna tiene maridos fijos imagínense como administradoras de un pueblo […]»; otro ejemplo, en la sentencia SRE-PSC-195/2021, se calificó como infractora la siguiente frase ubicada en *X*: «La gran diferencia entre la *pornstar* y la prostituta es que se está buscando una diputación… cobrará para que la vean en el canal del Congreso? (*sic*)».

---

[31] Véase sentencia: SRE-PSC-94/2022.

[32] Por citar un ejemplo, en la sentencia SUP-REP-160/2022, al analizar una nota periodística en la que a una candidata la identificaron como «verde e inexperta», «intolerante» y que no tenía «las tablas para lidiar», la Sala Superior expresamente señaló que, para que dichas frases fueran consideradas VPMG debían constituir estereotipos de género (par. 100) y que, en este caso no lo eran (par. 110).

En esta modalidad en la que se comete la VPMG, los hechos externos están constituidos por expresiones lingüísticas emitidas a través de frases, e incluso de «memes», y quien juzga debe valorar, a partir de un análisis al lenguaje, si se está en presencia o no de estereotipos de género. Para llegar a este punto, previamente debe estar demostrado, por ejemplo, que Pedro emitió cierta expresión en *Instagram* y que es el titular de la cuenta en la red social[33]. Una vez acreditados estos extremos fácticos, las expresiones vertidas se convierten en el objeto de la valoración. Llegado a este momento, identificar si las expresiones constituyen o no estereotipos de género, es en una tarea valorativa[34] de quien juzga —convirtiéndose en un problema de calificación[35] y no de prueba—, cuyo juicio de valor debe estar sustentado en los estereotipos que se suelen encontrar en una sociedad determinada.

Se trata de una tarea compleja mediante la cual el Tribunal debe analizar el lenguaje a partir de expresiones que van desde aspectos generales de la cultura popular mexicana a cuestiones para las que se requiere tener cierto conocimiento sobre la manera en que se emplean estas palabras en un territorio en específico. Al respecto, la autoridad jurisdiccional ha hecho un esfuerzo por marcar supuestas «pautas objetivas» en el análisis del lenguaje, ya que en la jurisprudencia 22/2024[36] estableció una metodología de estudio en la que fijó como parámetros: verificar el contexto, precisar la expresión, identificar la semántica, considerar los usos, costumbres y regionalismos en el lenguaje y analizar la intención.

Sin embargo, aunque la autoridad jurisdiccional ha realizado esfuerzos por reducir el margen de subjetividad en el análisis del lenguaje, las complejidades de esta tarea distan de estar resueltas. Un ejemplo reciente es la sentencia SRE-PSC-94/2024[37], en la que se analizó la siguiente frase publicada en *X* por una ciudadana en referencia a una diputada federal: «Así estaría el berrinche de DATO PROTEGIDO para que incluyeran a su esposa, que tuvieron que desmadrar las fórmulas para darle una candidatura. Cero pruebas y cero dudas». El asunto adquirió relevancia mediática al haberse calificado dicha expresión como VPMG, a pesar de que se argumentó que se trataba de una crítica al nepotismo y no de un mensaje sustentado en estereotipos de género.

---

[33] Muchos de los problemas de prueba en casos de VPMG también radican en demostrar la titularidad de determinadas personas en las redes sociales, por citar un ejemplo, véase la sentencia: SRE-PSC-45/2022.

[34] La identificación valorativa de un hecho debe realizarse, siguiendo a Michele Taruffo (2020, p. 128) cuando se está frente a normas que vinculan consecuencias jurídicas a hechos, que postulan algo más que la pura y simple existencia de un hecho en el sentido del propio término. Postulan, la existencia de hechos, comportamientos o estados, pero vinculan la relevancia jurídica al resultado de la valoración.

[35] Entendiéndose por problemas de calificación aquellos en los que la duda surge sobre si un determinado hecho, que no se discute, cae o no bajo el campo de aplicación de un determinado concepto contenido en el supuesto de hecho o en la consecuencia jurídica de una norma (Atienza, 2013, p. 432).

[36] De rubro: «Estereotipos de género en el lenguaje. Metodología para su análisis».

[37] Sentencia emitida por la Sala Especializada y confirmada por la Sala Superior en la diversa SUP-REP-401/2024 y acumulados.

Para identificar el sentido de las palabras, en ocasiones se han realizado diligencias de investigación o también se suele invocar hechos notorios. Para citar algunos ejemplos, en el expediente SRE-PSC-102/2021 una candidata perteneciente a una población indígena denunció una serie de publicaciones en *Facebook* en la que se realizaron diversas críticas a su candidatura, dentro de estas destacaba una que hacía alusión a que: «es tiempo de chicatanas», lo que en concepto de la denunciante estaba vinculado a su condición de mujer. En ese sentido, para determinar si la frase constituía o no un estereotipo de género era necesario saber cuál es el significado y en qué sentido se utiliza dicha palabra en la región mixteca de Oaxaca, lo que escapaba del conocimiento de las personas que participaban en la elaboración y resolución de la sentencia. En ese caso se realizó una solicitud de información a la Secretaría de Pueblos Indígenas y Afroamexicano de Oaxaca. No obstante, esto no resolvió el problema, porque la Secretaría informó al menos tres sentidos en los que podía utilizarse esta palabra y ninguno de ellos daba claridad sobre el sentido en que la empleó el emisor del mensaje.

En cuanto a los hechos notorios, en la sentencia SRE-PSC-41/2022, entre diversas imágenes editadas y frases en *Facebook* en contra de una candidata, se estudió la frase: «Su secretaria particular aprovechó y se subió al tapanco, sin ninguna experiencia política ni cargos importantes, la juanita se trepó, todo apunta a que […] impuso a su candidata». En una primera impresión, el apelativo «juanita» no tiene algún contenido que pueda considerarse alguna visión estereotipada. Sin embargo, en la sentencia se señaló que: «constituye un hecho notorio para esta autoridad jurisdiccional que la palabra juanita se realiza (*sic*) de una manera despectiva para referir que la denunciante será utilizada y en realidad no hará uso de sus facultades si llega a ganar. De esta manera se pone en duda su capacidad y su trayectoria, únicamente utilizando como referentes descalificativos y estereotipos establecidos por la sociedad».

Para justificar este hecho notorio, la autoridad jurisdiccional explicitó que dicho apelativo se originó en 2009, cuando un candidato electo conocido como «Juanito» pidió licencia el día que rindió protesta y cedió su lugar a una mujer. Posteriormente, el apelativo se les asignó a las mujeres en el proceso electoral de ese mismo año, pues con la finalidad de cubrir las cuotas de género, los partidos políticos designaron a diversas mujeres como candidatas, y una vez que fueron electas, cedieron su lugar a los suplentes, quienes eran hombres. Este hecho notorio puede encontrar sustento en diversos trabajos académicos y en sentencias mismas[38], ya que es el antecedente del principio de suplencia en las reglas de paridad. En este supuesto, dicha notoriedad puede justificarse de manera razonable. Empero, el Tribunal también ha invocado hechos notorios sobre cuestiones que son más bien de índole cotidiano, por ejemplo, en esa misma sentencia también se analizó y se invocó como hecho notorio el significado de palabras como «chucky», «novia de chucky» y «chimoltrufia».

[38] Véase sentencia: SUP-JDC-12624/2011.

Respecto a la regulación de los hechos notorios en el PES, la Ley Electoral prevé que estos no serán objeto de prueba (artículo 461, numeral 1 de la Ley Electoral). Entonces, en determinados casos, son utilizados para dotar de contenido a las palabras y así poder valorar si constituyen estereotipos o no. Así, los hechos notorios pueden ayudar a esos efectos; empero, esto no convierte al análisis del lenguaje en una labor libre de complejidades. En primer lugar, porque la mayoría de estas frases se realizan en las redes sociales, donde fluye una inmensa cantidad de información, e interactúan múltiples personas usuarias de distintas realidades, por lo que es difícil encontrar el sentido connotativo de las palabras que ahí se expresan. En segundo lugar, porque dado el diseño procesal del PES, cuando el órgano jurisdiccional invoca hechos notorios, no existe posibilidad de contradicción para las partes.

Sobre los hechos notorios, Anderson, Schum y Twining (2015, p. 336) sostienen que la idea de reconocer un consenso cognitivo que comprensa un acervo de conocimientos o creencias comunes resulta problemático. En igual sentido, Carmen Vázquez[39] (2022, pp. 145 y ss.), enfatiza en el problema que supone el consenso cognitivo de los hechos notorios, dadas las distintas clases o culturas; y enuncia diversas dificultades —de las cuáles tendría que hacerse cargo el órgano jurisdiccional— entre las que se encuentra cuán general debe ser el conocimiento del hecho en la sociedad de referencia.

En este punto, podría plantearse cuán general es el conocimiento de la sociedad respecto a que el apelativo «juanita» alude a una mujer que renuncia a su candidatura para cederla a un hombre; que «chucky» es un personaje identificado con diversas películas de terror, cuya trama ficticia hace referencia a un muñeco que es poseído por un asesino serial; o que la «chimoltrufia» es un personaje de la cultura popular mexicana, caracterizado como una «ama de casa» (*sic*) que es ridiculizada. También podría plantearse, desde el punto de vista lingüístico, si al emplear la palabra «chimoltrufia» como signo, una persona es capaz de asociarla mentalmente con el sentido que es utilizado en la cultura popular mexicana[40].

Para concluir, al intentar responder la pregunta: ¿cuál es la línea que permite distinguir entre una expresión que constituye un estereotipo de género y una que no lo es?, lo expuesto sugiere que tal línea no existe de forma clara u objetiva. Esto es así porque la calificación de las expresiones depende de una tarea valorativa basada en el contexto, los usos lingüísticos y los significados socialmente atribuidos a las palabras;

---

[39] La autora en este punto analiza la tesis P./J.74/2006, de rubro: «Hechos notorios. Conceptos general y jurídico», que es medularmente coincidente con las *Federal Rules of Eivdence*, particularmente la regla 201, el cual prevé que un hecho notorio no es susceptible de disputa razonable debido a que es ampliamente conocido dentro del ámbito territorial del tribunal del juicio; o puede ser determinado de forma acertada y fácilmente de fuentes cuya fiabilidad no puede ser razonablemente cuestionada.

[40] Al analizar el problema de los signos y la semiótica, Vitor de Paula Ramos (2023, p. 75) expone como ejemplo el uso de la palabra «unicornio» que, aunque este no exista, pese a no existir existe una idea de lo que representa, de modo que los signos propician asociaciones mentales, basadas en hábitos culturales adquiridos.

elementos que no admiten una determinación unívoca y que varían en el entorno social y cultural en que se emiten.

Así, a pesar de que desde el punto de vista del razonamiento probatorio las autoridades que intervienen en el proceso tienen la posibilidad de recabar elementos de juicio con la finalidad de dotar de contenido el significado de las palabras, o bien acudir a hechos notorios debidamente justificados, siempre existirá un margen irreductible de subjetividad. Sin embargo, esta decisión no debe quedar al libre arbitrio de la persona juzgadora; máxime, cuando se está en confrontación con otro derecho fundamental: la libertad de expresión.

En ese sentido, el control de la decisión debe articularse a partir de la exigencia de justificación explícita, en la que se identifiquen las inferencias realizadas, las máximas de la experiencia invocadas o los elementos que permiten concluir que una palabra tiene cierto significado, así como las razones por las cuales una determinada expresión es calificada como estereotipada. Esto, con la finalidad de que la decisión sea susceptible de control externo y evitar que el análisis al lenguaje se convierta en un ejercicio discrecional incompatible con las exigencias del razonamiento probatorio.

3. ANÁLISIS DE LOS CRITERIOS PROBATORIOS DEL TRIBUNAL ELECTORAL

Como mencioné en el apartado introductorio, el Tribunal Electoral ha establecido diversos criterios que tienen como propósito atender los problemas de prueba que se presentan al momento de resolver denuncias sobre VPMG, entre los cuales destacan los siguientes[41]:

— *Inversión de la carga de la prueba*. En casos de VPMG, el Tribunal ha establecido que debe operar una excepción a la regla general sobre el onus probandi. Corresponde a la persona denunciada desvirtuar de manera fehaciente los hechos, especialmente cuando se trata de actos ocurridos en espacios privados y sin presencia de testigos, de modo que no se exija un estándar imposible de prueba a la víctima.

— *Valor de las declaraciones de la denunciante.* La prueba aportada por la presunta víctima goza de presunción de veracidad y constituye un elemento fundamental para acreditar los hechos. Su testimonio puede tener valor probatorio pleno, especialmente cuando se vincula con otros indicios, y debe ser evaluado con perspectiva de género, evitando exigencias probatorias estereotipadas que obstaculicen su acceso a la justicia.

Ambos criterios, según ha sostenido el Tribunal Electoral, tienen como fundamento la aplicación de la perspectiva de género como método de análisis para juzgar.

[41] Los extractos citados corresponden a la sentencia SUP-REC-91/2020, que fue la primera en establecer el criterio de reversión de la carga de la prueba.

Por ese motivo, en principio, se expondrá, en forma descriptiva, algunos postulados sobre lo que implica juzgar con este enfoque y, posteriormente, se analizarán los referidos criterios probatorios.

3.1. Juzgar con perspectiva de género como método de análisis

Hablar sobre perspectiva de género y su relación con la prueba es un tema más que complejo, por ese motivo, se anticipa que este trabajo no tiene ánimo de exhaustividad respecto a todas las aristas que pueden comprender estas dos materias y su relación. La pretensión es únicamente exponer un breve panorama sobre lo que implica juzgar con perspectiva de género, con el objeto de confrontar ese marco conceptual con la forma en que el Tribunal Electoral ha interpretado y aplicado dicho enfoque al momento de justificar los criterios probatorios que emplea en la resolución de los casos de VPMG.

También es necesario precisar que *juzgar* con perspectiva de género es solamente uno de sus enfoques, pues esta herramienta permea en cuestiones que van desde lo cotidiano, como adoptar un lenguaje inclusivo; lo gubernamental, como generar políticas públicas; lo legislativo, es decir, crear leyes con esta visión, etcétera.

Con estas salvedades se tomará como referente la idea de juzgar con perspectiva de género que busca cuestionar la aparente neutralidad con la que están construidas las normas y el Derecho[42]. En palabras de Francisca Pou (2014, pp. 125-128) adoptar este enfoque implica hacerse cargo de que las personas sufren algún tipo de injusticia, opresión o desventaja por motivos de sexo o género que requiere aprender a manejar las herramientas analíticas y argumentales necesarias para administrar justicia. Para la constitucionalista, esta visión exige apartarse de la idea de entender a las partes como iguales, y reconocer la existencia de inequidades de género. En la misma lógica, Olga Fuentes Soriano (2022, p. 73), precisa que esta herramienta aspira a que los parámetros que utiliza el sistema de justicia para interpretar y aplicar la ley no refuercen a través de una neutralidad axiológica vinculada a la igualdad formal, las relaciones de poder de los hombres sobre las mujeres, consolidando, de esta manera, la discriminación de estas últimas.

En forma más específica en su relación con la prueba, Marcela Araya (2020, pp. 38 y ss.)[43] analiza el papel que debería tener la perspectiva de género en los momentos de la actividad probatoria propuestos por Jordi Ferrer[44]. En cuanto a la conformación del acervo de prueba, la autora precisa que la dificultad se encuentra en integrar un conjunto rico y variado de elementos de juicio que haga factible la condena de la persona materialmente culpable, y destaca en la importancia que su-

[42] Un análisis muy amplio sobre esto puede verse en: Mahoney (1997).
[43] Se precisa que la autora sitúa su análisis específicamente a la materia penal.
[44] Sobre los momentos de la actividad probatoria véase: Ferrer, 2007, pp. 41-47.

pone, entre otras, el evitar interferir provocando discriminación dada la existencia de prejuicios. En esa misma línea, la Suprema Corte de Justicia de la Nación[45] ha establecido como una de las pautas obligatorias para juzgar con perspectiva de género que en caso de que el material probatorio no sea suficiente para aclarar la situación de violencia, vulnerabilidad o discriminación por razones de género, se deben ordenar las pruebas necesarias para visibilizar dichas situaciones.

Resulta interesante que, a pesar de que existen posturas y puntos divergentes sobre el papel de la perspectiva de género en cada uno de los momentos de la actividad probatoria, hay una medular coincidencia en que la conformación de elementos de juicio es un punto clave. Por ejemplo, la autora citada sostiene que enfocar los esfuerzos en esta etapa rendirá frutos en la valoración probatoria (Araya, 2020, p. 64). En la misma lógica José Luis Ramírez Ortiz (2020, p. 244) quien es crítico sobre la utilidad epistémica de la perspectiva de género —específicamente hablando de la prueba testimonial como única prueba de cargo— concluye que el deber de los Estados de llevar a cabo investigaciones diligentes y efectivas para esclarecer los hechos en cuestión es el mejor camino para combatir el fenómeno de la violencia de género. Igualmente, Jordi Ferrer (2019[46]) quien realiza comentarios sobre la perspectiva de género en el segundo momento de la actividad probatoria —tema sobre el que se volverá más adelante—, refiere que la autoridad que investiga debe ser consciente de las situaciones de dominación por razón de género y ello exigirá un papel proactivo en la búsqueda de elementos de juicio para demostrar los hechos.

Siguiendo el camino que propone Marcela Araya (2020, p. 40) resalta la importancia de la búsqueda de prueba de contexto que permita a quienes investigan ampliar los márgenes de los elementos de juicio y poner atención a las evidencias que sean relevantes, incluyendo todas las circunstancias que rodean la denuncia. Sobre este punto, Sandra Serrano (2019, p. 26) precisa que este análisis implica analizar el contexto de la persona que denuncia en forma particular, pero que esto no basta, sino que es necesario entender también el contexto del sector en opresión al que pertenece a fin de hacernos cargo de la desigualdad estructural como lo exige la perspectiva de género. Ambas autoras coinciden en que dichas pautas se marcaron por primera ocasión en el caso Campo Algodonero *vs*. México de la Corte Interamericana de Derechos Humanos.

En cuanto al momento de la valoración probatoria Araya (2020, p. 47) destaca el papel de las máximas de la experiencia y afirma que debe asumirse con mucho rigor el empleo de generalizaciones empíricas depuradas y descartar aquellas que encubran meros prejuicios, arquetipos y preconcepciones sobre el rol ideal de las mujeres en la sociedad. En este caso, como expresé en párrafos arriba, el reto se encuentra no solo

---

[45] Jurisprudencia 1a./J. 22/2016 (10a.), de rubro: «Acceso a la justicia en condiciones de igualdad. Elementos para juzgar con perspectiva de género».

[46] Conferencia impartida por Jordi Ferrer Beltrán el 26 de noviembre de 2019, organizada por la Dirección General de Derechos Humanos de la Suprema Corte de Justicia de la Nación, disponible en: *https://www.youtube.com/watch?v=SlsoI3WQLy4*.

al evitar los estereotipos, sino al identificarlos como la razón de la acción u omisión. Por su parte, el Alto Tribunal mexicano, en la jurisprudencia citada, ha enfatizado en que se deben cuestionar los hechos y valorar las pruebas desechando cualquier estereotipo o prejuicio de género. Al respecto, Jordi Ferrer (2007, pp. 47 y 91) ha sostenido que en este momento *reina sin competencia el valor de la averiguación de la verdad*[47], en el cual se evalúa el apoyo empírico que los elementos de juicio aportan individual y conjuntamente a las diversas hipótesis fácticas disponibles sobre lo ocurrido. Con esta base, el profesor citado (2019) ha sostenido que valorar la prueba con perspectiva de género no es una situación distinta que hacer una valoración racional.

En cuanto a la decisión sobre los hechos probados, Marcela Araya concluye que sería posible operar con estándares de prueba diferenciados, con el fin de aceptar una hipótesis como probada en ciertos tipos de delitos que exhiben dificultades probatorias, como ocurre en los casos de violencia contra la mujer por razones de género (2020, p. 62). Del recorrido que realiza, se desprende que la autora coincide con la propuesta de Ferrer Beltrán (2021, p. 149) en cuanto a que los estándares de prueba, al ser una decisión política, deben ser definidos por el órgano legislativo. Sostiene que, a su parecer, frente a las propuestas que se formulan con desmedido entusiasmo en torno a rebajar el estándar de prueba —refiriéndose a los delitos de naturaleza patriarcal—, no es posible *prima facie* responder inflexiblemente negando esa posibilidad; no obstante, precisa que la aceptación de aquella flexibilización o rebaja debe pasar necesariamente por una reforma legal que recoja la decisión soberana del reparto del error (Araya, 2020, p. 61).

Una vez hecho este breve recorrido, es posible desprender que tanto en la doctrina como en la jurisprudencia que se ha analizado existe un énfasis relevante en que las autoridades que intervienen en la instrucción y resolución de los casos asuman que el Derecho y su aplicación se encuentran influidos por estereotipos, prejuicios y concepciones erróneas sobre el papel que históricamente se ha asignado a las mujeres en los distintos ámbitos. En ese contexto, el análisis de las sentencias del Tribunal Electoral permite identifica un esfuerzo por atender dicho imperativo, aunque, en el intento por cumplirlo, en ocasiones se adoptan concepciones problemáticas sobre la tarea que supone juzgar con perspectiva de género; cuestión que se examina en los apartados siguientes.

Otra de las tareas más importantes está en el primer momento de la actividad probatoria, es decir en la conformación del conjunto de elementos de juicio. En este caso, dado que el PES funciona por medio de un sistema dual corresponde a la autoridad administrativa electoral[48] realizar la investigación de las denuncias y, en su

[47] Jordi Ferrer (2007, p. 91) precisa que nunca un conjunto de elementos de juicio, por grande y relevante que este sea, permitirá tener certezas racionales sobre la verdad de una hipótesis.

[48] En específico a nivel federal la autoridad encargada de investigar las denuncias de VPMG es el Instituto Nacional Electoral, mientras que a nivel estatal investigan los Organismos Públicos Electorales Locales de cada entidad.

caso, una vez que el expediente llega a la autoridad jurisdiccional esta tiene la facultad probatoria de devolverlo a la autoridad administrativa[49] y señalar de forma específica las diligencias que debe realizar para allegarse de las pruebas necesarias.

Para concluir este apartado, es importante mencionar que los postulados de la perspectiva de género que, sin ánimo de exhaustividad, han sido expuestos, no buscan subsanar la insuficiencia probatoria y tampoco ir en contra de normas epistémicas, por lo que tampoco se pretende algo contrario o distinto a realizar una valoración racional de la prueba.

Ahora bien, la importancia de hablar en términos específicos sobre perspectiva de género estriba, al menos, en dos cuestiones. La primera, como una manera de reivindicación que busca cuestionar las bases con las que históricamente se crearon las instituciones y las leyes, así como los patrones androcéntricos que han generado relaciones de poder[50]. La segunda, porque los estereotipos, sesgos y prejuicios de género hacia las mujeres se encuentran —lamentablemente— inmersos en la sociedad y son asumidos como normales tan es así que en ocasiones no se opera en el orden de las intenciones conscientes[51], por lo que es importante e indispensable poner un énfasis específico para identificarlos.

3.2. Inversión de la carga de la prueba

La Sala Superior del Tribunal ha establecido lo siguiente: «la reversión de las cargas probatorias opera a favor de la víctima en casos de violencia política en razón de género ante situaciones de dificultad probatoria, por lo que la persona denunciada como responsable tendrá la carga reforzada de desvirtuar de manera fehaciente los hechos de violencia que se le atribuyen en la denuncia»[52]. El fundamento de este criterio es la dificultad de las personas denunciadas para evidenciar la violencia, haciendo hincapié en situaciones de discriminación y desigualdad, así como la aplicación de la perspectiva de género. El Tribunal ha puesto el acento en que se debe evitar trasladar a las posibles víctimas la responsabilidad de acreditar los hechos en casos de VPMG, en aras de no obstaculizar el acceso a la justicia[53].

Cabe apuntar que el Tribunal ha matizado su postura, por ejemplo, en la sentencia SUP-JE-122/2022 pasó de decir que la regla habitual era invertir la carga de

---

[49] De conformidad con el artículo 476, numeral 2, inciso b de la Ley Electoral.

[50] Sobre este punto, véase: Fuentes (2022, p. 68).

[51] Al respecto, Pierre Bourdieu, en su texto *Symbolic violence* (1996, pp. 1-4) sostiene que la violencia contra las mujeres está inmersa en términos simbólicos, de tal forma que en ciertos casos no se opera en el orden de las intenciones conscientes, dado que la violencia atiende a una estructura de dominación fuera de todo consentimiento voluntario, consciente, y de toda coerción directamente ejercida.

[52] Jurisprudencia 8/2023, de rubro: «Reversión de la carga probatoria. Procede en casos de violencia política en razón de género a favor de la víctima ante la constatación de dificultades probatorias».

[53] Véase sentencia SUP-REC-133/2020.

la prueba, a sostener que esto no es así en todos los casos, sino que debía atender al contexto y particularidades del supuesto concreto. Por su parte en la determinación SUP-REC-200/2022 estableció que la inversión de la carga de la prueba debe ser comunicada a quien se denuncia, ya que de lo contrario no existe otra manera para hacerle saber que le recae la carga de desvirtuar los hechos imputados, de forma tal que se le debe informar que su silencio o inactividad procesal tendrían como posible consecuencia que se le considere como persona infractora.

Sobre la carga de la prueba, Jordi Ferrer (2019, p. 58) refiere que es una institución probatoria residual o subsidiaria, que operaría solo ante el fracaso de la actividad probatoria de las partes conducente a acreditar los hechos del caso, con la finalidad de ofrecer un criterio a quien juzga para dar cumplimiento a su obligación de resolver todos los casos[54]. Por su parte, Michele Taruffo (2010, p. 255) indica que es una regla de juicio consistente en que quien no demuestre la verdad de los hechos que tiene la carga de probar será derrotado, pues quien juzga debe establecer las consecuencias de la falta de prueba de los hechos que una parte ha alegado, decidiendo en su contra.

En forma específica respecto a la inversión o carga dinámica Michele Taruffo (2010, p. 258) explica que lo que se busca con dicha figura es facilitar la tutela de derechos de personas que se consideran débiles en el proceso y que encontrarían graves problemas, o la imposibilidad de proporcionar la prueba sobre los hechos en que se fundan sus pretensiones. El autor realiza diversas críticas a esta postura, por ejemplo, que cuando esto se realiza en sede jurisdiccional —situación distinta a hacerlo en sede legislativa mediante presunciones legales— quien juzga está realizando reglas de juicio distintas a las que la Ley prevé, atribuyéndose un poder discrecional que ninguna norma les ha conferido, además de que dicha decisión tiene como base sus preferencias subjetivas de la solución justa de la controversia a un caso concreto (Taruffo, 2010, pp. 262-263).

El autor (Taruffo, 2010, p. 263) sostiene igualmente que le parece débil la justificación que se funda en la mayor proximidad de una parte a las pruebas; a su parecer, no es necesario invertir la distribución de cargas probatorias, puesto que, sería suficiente y más correcto prever una orden de exhibición efectiva y adecuadamente sancionada, que impusiera a la parte que dispone de la prueba que sería útil a la otra parte aportarla al juicio.

---

[54] El autor citado precisa que la carga de la prueba tiene dos dimensiones: «objetiva y subjetiva, de las que la primera sería prioritaria y la segunda una consecuencia derivada. La carga de la prueba en su dimensión objetiva es una prescripción dirigida al juez, indicándole el sentido que debe dar a la resolución judicial en caso de ausencia de prueba suficiente sobre los hechos. Se trata pues de una regla de aplicación subsidiaria en el momento de la toma de decisión. La carga de la prueba en su dimensión subjetiva, en cambio, se dirige a las partes, indicando qué hechos deben probar si quieren obtener un resultado favorable en el proceso». Sin embargo, respecto a estas dimensiones, a su parecer, son facetas totalmente distintas, dado que van dirigidas a diferentes partes y regulan actividades probatorias diferentes, por lo que no hay ninguna razón para sostener que la carga subjetiva es consecuencia de la objetiva (Ferrer, 2019, p. 69).

Por otro lado, De Paula Ramos (2020, p. 112) —hablando específicamente del proceso civil— ha enfatizado en que cualquier proceso que tenga pretensiones de justicia debe evitar basarse en la regla de juicio de la «carga», a partir de la asunción de que una sentencia basada en «hechos» que no ocurrieron es una sentencia equivocada. El autor considera que «las *cargas* sin consecuencias concretas vinculadas a la (in)actividad de la parte no son un verdadero estimulo jurídico para que la parte practique una prueba cuando esta le es desfavorable» (2020, p. 117). En ese sentido, su propuesta es transitar de la «carga», al «deber de probar», a partir de una potestad coactiva prevista por el órgano legislativo con la finalidad de que la persona juzgadora estimule a las partes para considerar pruebas que se estimen relevantes para el proceso (2020, p. 118).

En igual sentido, Jordi Ferrer (2019, pp. 79 y ss.) aduce que existen otras reglas capaces de generar incentivos a las partes para aportar pruebas al proceso. A su juicio, este cometido no se cumple invirtiendo la carga de la prueba; precisa que existe una confusión —derivada de lo que a su parecer es una aparente o mal entendida relación entre la carga objetiva y subjetiva de la prueba— entre el deber de aportar pruebas sobre la base del principio de colaboración procesal que conlleva a la obligación de realizar un acto, con imponer la carga de la prueba.

Ahora bien, el fundamento que utiliza el Tribunal para la inversión de la carga de la prueba en casos de VPMG es la asunción de una posición de desventaja entre las partes; pues, se asume que la persona denunciada no solo está en una mejor posición para probar, sino estructuralmente hablando respecto a la presunta víctima, de tal suerte busca materializar el principio de igualdad entre las partes. Al respecto, en casos de VPMG esto no siempre es así. Desde luego pueden presentarse supuestos en los cuales la persona que se denuncie esté en mejores condiciones para probar, pero no es una máxima absoluta.

Por ejemplo, piénsese en un asunto en el cual quien denuncia es una Diputada que puede desde el poder público allegarse de los elementos de juicio necesarios frente a un ciudadano o ciudadana común; o en situaciones donde incluso la denunciante puede estar en mejores condiciones de probar, como sería el posible caso de la Consejera Presidenta de un Organismo Electoral Local que ocupa el más alto cargo en el órgano de dirección superior, frente a cualquier otra persona funcionaria pública de ese Organismo. También puede haber situaciones entre pares, por ejemplo, dos funcionarias que ocupan un mismo cargo en un órgano colegiado. No obstante, matizar el criterio como lo ha hecho la Sala Superior del Tribunal no resuelve los problemas, pues nunca existirá certeza para la parte denunciada respecto cuando procederá la inversión.

Otro punto importante que debe tener en cuenta la autoridad electoral es la dificultad para demostrar hechos negativos. Pongo como ejemplo las circunstancias fácticas de la sentencia SUP-REC-2214/2021 y acumulados. En este caso, la candidata a la presidencia municipal del Ayuntamiento de Atlautla, Estado de Méxi-

co [55] y el partido político que la postuló, interpusieron un medio de impugnación en contra de la declaración de validez de la elección, ya que durante el desarrollo de la campaña política vandalizaron diversas bardas con propaganda electoral de su candidatura con una serie de insultos y amenazas. La autoridad jurisdiccional concluyó que, si bien no se demostró la autoría de las pintas de bardas, el estudio de la VPMG con perspectiva de género llevaba a la conclusión de que «exigir demostrar la autoría de los actos implicaría imponer un estándar de prueba prácticamente imposible de superar».

Cabe señalar que en este caso el órgano jurisdiccional empleó de manera imprecisa la expresión «estándar de prueba», en tanto que no se refiere al umbral de suficiencia probatoria, sino a la fijación del objeto de la prueba, en concreto, a la exigencia de acreditar la autoría de los hechos denunciados. Asimismo, se precisa que en este asunto el Tribunal no mencionó de manera explícita que estuviera invirtiendo la carga de la prueba, sino que argumentó que el instituto político de la planilla ganadora no demostró haberse deslindado de las conductas, es decir, como una especie de culpa *in vigilando*. Se puntualiza también que la consecuencia no fue estimar la comisión de VPMG a alguna persona en específico, pero sí la nulidad de la elección del municipio de Atlautla, y dicha determinación sí afectó a la planilla que había resultado ganadora.

A pesar de esto, el caso es útil para ilustrar las dificultades que representa acreditar estos hechos negativos, en el supuesto concreto, ¿de qué manera podría una persona o un partido político demostrar que no pintó y no mandó pintar las bardas?

De igual forma, resulta discutible invertir la carga de la prueba tomando como sustento que el juzgar con perspectiva de género implica no trasladar a la denunciante la responsabilidad de aportar pruebas, puesto que en todo caso ello podría traducirse más bien en otros incentivos para producir pruebas relevantes, como no desechar la denuncia por falta de pruebas (artículo 471, numeral 5, inciso c) de la Ley Electoral) y realizar las diligencias de investigación que se estimen pertinentes para allegarse de todos los elementos de juicio posibles; o bien, en última instancia, generar el incentivo de aportar determinadas pruebas al proceso con la potestad que tiene la autoridad electoral de imponer medidas de apremio (artículo 461, numeral 10 de la Ley Electoral). Agotando estas posibilidades, y en caso de no haberse cumplido el nivel de corroboración requerido, al tratarse de un procedimiento sancionador al que le son aplicables las garantías del *ius puniendi* del derecho penal, en el que se encuentra la presunción de inocencia como regla de juicio, la manera en que debería de resolverse es determinando la inexistencia de la comisión de VPMG.

Finalmente, parte de la insistencia del Tribunal de sostener esta figura es, en ocasiones, la imprecisión conceptual. Por ejemplo, en la sentencia SUP-REP-245/2020

[55] Un asunto similar puede verse en la sentencia: SUP-REC-1861/2021, en ese caso se anuló la elección de Iliatenco, Guerrero.

y acumulados, la Sala Superior dispuso que: «respecto al estándar probatorio para configurar dicha conducta, este órgano jurisdiccional ha determinado que es insuficiente la declaración de la inversión de las cargas probatorias, pues deben tenerse elementos probatorios que conduzcan a tener por acreditada una infracción». Estas líneas carecen de sentido si se toma como referente que el estándar tiene la función de determinar el umbral de suficiencia probatoria a partir del cual una hipótesis sobre los hechos deberá considerarse probada (Ferrer, 2021, p. 33) y que la carga de la prueba es una figura que opera en forma subsidiaria cuando este estándar no se cumplió (Ferrer, 2019, pp. 58-68), del tal forma que se invierte justamente porque no existen en ese momento suficientes elementos probatorios sobre los hechos del caso.

Como se mencionó al inicio de este apartado la Sala Superior del Tribunal comenzó sosteniendo en forma tajante que en casos de VPGM opera la figura de la inversión de la carga de la prueba, criterio que ha ido matizando. Sin embargo, los criterios referidos siguen replicándose en sus sentencias, muchas veces a partir de apreciaciones incorrectas sobre lo que implica invertir la carga de la prueba. Además, los matices que ha realizado únicamente generan falta de certeza jurídica para las partes en el proceso. Vale la pena señalar que, al ser un criterio de la Sala Superior, por la manera en que opera el sistema de precedentes electorales, este debe ser aplicado no solo en forma vertical en sus propias sentencias, sino que tiene un muy importante impacto horizontal en las Salas Regionales, los Tribunales locales e inclusive en el Instituto Nacional Electoral y en los Institutos locales cuando ejercen funciones materialmente jurisdiccionales, como sería el caso de las medidas cautelares en el PES. Por las razones expuestas es que concluyo que el Tribunal haría bien en replantear este criterio.

3.3. Consideraciones en torno a la declaración de quien denuncia VPMG

La Sala Superior ha reiterado que los actos de violencia por razones de género suelen tener lugar en espacios privados donde solo se encuentran la víctima y la persona agresora. Por este motivo ha resaltado que los hechos no pueden someterse a «un estándar de imposible prueba», y que la prueba que aporta la denunciante es pieza fundamental sobre el hecho. Asimismo, ha enfatizado en que las manifestaciones de la presunta víctima en conjunto con «otro tipo de indicios» pueden construir una «prueba circunstancial»[56] con «valor pleno»[57].

---

[56] Al respecto, si se toma como referente que todo argumento probatorio es una cadena de razonamiento conformada por lo que Anderson, Schum y Twining (2015, pp. 94-96) denominan *probanda* —proposiciones a ser probadas—, las cuales deben ir de las *probanda penúltimas* —proposiciones simples— a un *probandum final* y que en su concepto rara vez un probandum final es una proposición simple, entonces se puede decir que en realidad toda prueba producto de una cadena de inferencias y, en consecuencia, circunstancial, por lo que en realidad esto no es una excepción para los casos de VPMG como lo refiere el Tribunal

[57] Como se verá más adelante, desde el punto de vista epistemológico, es impreciso hablar de una prueba con «valor pleno».

En cuanto a este tópico, lo primero que debe aclararse es que cuando el órgano jurisdiccional hace alusión a las manifestaciones de la presunta víctima se refiere a lo que la denunciante expresa en el escrito de denuncia o en la audiencia de pruebas y alegatos, pues se reitera que el PES no contempla pruebas personales. Por otro lado, de la revisión a las sentencias de la materia, se advierte que el Tribunal no ha considerado la comisión de VPMG únicamente con la sola manifestación de quien denuncia, pero lo que sí ha hecho es que cuando se hace valer que las conductas ocurrieron es espacios cerrados, el Tribunal toma por acreditadas las manifestaciones presuntamente dichas y verifica otras circunstancias que rodearon el caso[58].

En este aspecto la preocupación de la autoridad jurisdiccional estriba en la dificultad de demostrar la comisión de VPMG y en la manera de eliminar esta incertidumbre. Es posible que el Tribunal inspire sus criterios en las dificultades probatorias de los delitos contra la libertad sexual del Derecho penal. Esto puede ser así porque, entre otras determinaciones, la Primera Sala de la Suprema Corte de Justicia de la Nación en el amparo directo en revisión 3186/2016, analizó la valoración del testimonio de la posible víctima de delito cuando la misma es la única prueba de cargo.

El Alto Tribunal especificó, en la parte que interesa, que: a) los delitos sexuales son un tipo de agresión que, en general, se producen en ausencia de otras personas más allá de la presunta víctima y la persona o personas agresoras, por lo que se requieren pruebas sobre otras conductas; y b) se debe analizar la declaración de la posible víctima en conjunto con otros elementos de convicción, recordando que la misma es prueba fundamental. Ambos postulados han sido en cierta medida trasladados por el Tribunal Electoral para los casos de VPMG como parte de lo que implica juzgar con perspectiva de género.

Sobre el valor probatorio del testimonio de quien denuncia tratándose de violencia sexual, la atención se encuentra en valorar los siguientes ítems: a) la credibilidad subjetiva de la presunta víctima; b) la verosimilitud del testimonio; y c) la persistencia en la incriminación. Así lo ha establecido la jurisprudencia española[59], cuyos elementos son medularmente coincidentes con los establecidos por la Corte mexicana en el amparo directo en revisión 3186/2016.

Empero, debe tomarse en cuenta que la VPMG no es una infracción que —por excelencia— sea de realización oculta, como sí lo son los delitos contra la libertad sexual, pues si bien las conductas no siempre serán necesariamente públicas o fáciles de demostrar, también es verdad que las mismas se suscitan en la arena político-electoral, así como en el ejercicio de cargos de elección popular, por lo que la autoridad instructora cuenta, al menos, con dicho margen contextual para llevar a cabo su

---

[58] Por citar algunos ejemplos, véase las sentencias: SUP-REP-477/2021 y SUP-REP-394/2021, ambas fueron resueltas en sentido de declarar la inexistencia de la comisión de VPMG a pesar de haberse tomado por ciertas las manifestaciones dichas en espacios privados.

[59] Sobre el análisis para valorar estos ítems, véase: Ramírez, 2020, p. 211; Fuentes, 2022, p. 75; Araya, 2020, pp. 52-53.

investigación. Por estos motivos, no en todos los contextos resulta adecuado aplicar exactamente el mismo tratamiento que en los delitos de naturaleza sexual, así como tampoco asumir que el debate en torno al testimonio como única prueba de cargo y su forma de corroboración sea necesariamente equivalente[60].

De esta manera, tomar por ciertas las manifestaciones de quien denuncia sin desplegar actividades de recolección de pruebas cuando se aduce que las conductas ocurrieron espacios cerrados, puede no ser la mejor alternativa desde una perspectiva epistémica, sin que pase desapercibido que todo procedimiento dependerá del supuesto bajo examen y el contenido de la investigación. Aunado a ello, se puede generar un incentivo para que, en forma contraria a lo que busca la perspectiva de género, la autoridad instructora relaje las actividades indagatorias. En este punto, se reitera el consenso que existe sobre la importancia y utilidad de encaminar los esfuerzos en la conformación de los elementos de juicio (Ramírez, 2020, p. 244; Araya, 2020, p. 64; Ferrer, 2019).

Ahora bien, desde luego es innegable la importancia de la manifestación de la denunciante, pues a partir de esta la autoridad instructora debe dirigir sus indagatorias, también en este punto es crucial cumplir con el imperativo reiterado de la perspectiva de género de no partir de sesgos, prejuicios o estereotipos para omitir investigar o para conducir su investigación. En estos supuestos, las dificultades probatorias no estriban por excelencia en contar solo con la manifestación de quien denuncia, sino en que debe demostrarse que los actos u omisiones estén basados en *elementos de género* y esto nos lleva de regreso a lo dicho en apartados atrás, sobre el difícil razonamiento que va de los hechos externos a los internos, así como lo problemático de intentar demostrar el impacto y la afectación desproporcionada.

Por otro lado, si se asume la premisa reiterada consistente en que mientras más rico sea el conjunto de elementos de juicio del que dispongamos para tomar una decisión, habrá mayor probabilidad de acierto (Ferrer, 2019, p. 55 y 2021, p. 22), valdría la pena reflexionar en la propuesta de modificar la Ley Electoral para admitir otro tipo de pruebas al proceso y no solo la documental y la técnica. La autoridad administrativa electoral lo que suele hacer para subsanar esto es citar a personas involucradas en el caso a realizar una especie de declaración frente a un funcionario o funcionaria investida de fe pública —dicha figura se denomina «oficialía electoral» (artículo 60 bis, inciso d) de la Ley Electoral)— y admite estas pruebas como documentales de carácter público[61].

En ese sentido, los problemas sobre la regulación probatoria en el PES y la manera en que se busca subsanar imposibilidad de incluir en el proceso otro tipo de pruebas son, al menos, los siguientes:

---

[60] Véanse los debates de Ramírez, Arena, Casiraghi, Fuentes y Gama en torno a la prueba testimonial como única prueba de cargo en delitos de violencia sexual la revista *Quaestio Facti*, número 1, disponible en: *https://www.quaestiofacti.com/es/paginas/numeros*.

[61] Se puede ver en las sentencias: SRE-PSC-2/2021 —se precisa que esta sentencia de la Sala Regional Especializada que fue revocada por la Sala Superior— y SUP-REP-394/2021.

El primero, que el valor probatorio la prueba documental está regulada en el PES de manera tasada por lo que no existe un auténtico *razonamiento* al valorar la oficialía electoral (artículo 462, numeral 2 de la Ley Electoral). Esto es un problema porque se trata de un documento con contenido testimonial y aunque este sea realizado por una persona investida de fe pública con la ventaja de que el testimonio permanecerá en el tiempo, esta circunstancia no hace que el testimonio goce, en sí mismo, de un nivel superior respecto a su contenido, es decir, a su justificación epistémica, como lo sostiene Vitor de Paula Ramos (2023, p. 158). Por tal motivo, es equivocado que se hable de «prueba plena» al valorar una oficialía electoral, ya que siguiendo al autor referido (2023, p. 170), desde el punto de vista epistémico, todas las pruebas pueden, en abstracto, aumentar o disminuir la corroboración de las hipótesis acerca de los hechos, por lo que una única prueba puede aportar un valor muy alto o bajo.

El segundo problema es que no se garantiza el principio de contradicción, pues como refiere Vitor de Paula Ramos (2023, p. 35) el alcance de la contradicción respecto al documento aportado se suele limitar al propio documento y a su contenido, posibilitándose la prueba en contrario, de forma tal que, si la parte tiene la oportunidad de manifestarse sobre los documentos aportados, se habrá satisfecho el derecho de contradicción. Al respecto, el autor enfatiza en que se debe entender el contexto en el que se elabora el documento, su forma de creación, su relación causal con el mundo, el contexto en el que se produjo, etcétera, para que realmente pueda ser objeto de contradicción y futura valoración (2023, p. 168).

En todo caso, la propuesta de incluir otro tipo de pruebas al PES es una tarea que debería realizarse en sede legislativa y resulta complicado que tratándose de una materia en la que están en juego distintos intereses políticos, se ponga un auténtico interés a las cuestiones probatorias.

4. CONCLUSIONES

Es indudable la importancia de que el Estado garantice un papel proactivo de las mujeres en la vida política, el cual evidentemente debe ejercerse en forma libre de violencia y discriminación. Esta es una premisa innegable, tanto desde el punto de vista de una teoría de la democracia que pretenda hacer tangible el mandato constitucional que hace recaer la soberanía popular en el pueblo mediante elecciones libres y auténticas, como desde el punto de vista de los derechos humanos de las mujeres, principalmente el de la igualdad. Ahora bien, el modelo punitivo por el que ha optado el Estado presenta diversas complicaciones. La pretensión del presente trabajo fue enfocarse en forma específica de los problemas de prueba, con la intención de hacer notar que la manera en la que está regulada la VPMG en la legislación, así como la forma en que ha sido interpretada por la jurisprudencia, involucra ciertas complejidades. Asimismo, se ha buscado exponer las razones por las cuales resulta necesario replantear los criterios adoptados para intentar resolver estos problemas,

tanto en sede legislativa como jurisdiccional. En ese sentido, este estudio busca contribuir a la reflexión, el análisis y el debate en torno a las siguientes premisas:

1. En sede jurisdiccional:

1.1. Es fundamental que las personas juzgadoras eviten recurrir a máximas de la experiencia epistémicamente deficientes, ya sea al identificar estereotipos de género como causa de una acción u omisión, o al intentar excluir dichos estereotipos como herramientas inferenciales.

1.2. Reflexionar sobre la manera en que la afectación desproporcionada y el impacto diferenciado deben ser contextualizados, a partir del principio de igualdad en su dimensión material y de la noción de discriminación indirecta, evitando recurrir únicamente a premisas genéricas como justificación para la actualización de estos supuestos.

1.3. Replantear si es conveniente o no continuar haciendo un análisis en sede jurisdiccional sobre las razones de género cuando el órgano legislativo ya adscribió dichos motivos en los catálogos enunciativos de la Ley Electoral y la Ley de Acceso.

1.4. Reconocer la importancia de que la autoridad jurisdiccional justifique las razones por las cuales considera que determinada expresión lingüística constituye, o no, un estereotipo de género.

1.5. Plantear la posibilidad de apartarse del criterio que asocia la implementación de la perspectiva de género con la inversión de la carga de la prueba, y enfocar los esfuerzos en identificar otros incentivos para producir pruebas relevantes.

1.6. Replantear la conveniencia o no de asumir como ciertas las declaraciones de quien denuncia VPMG cuando se argumenta que las conductas ocurrieron en espacios privados, y enfocar los esfuerzos en la investigación de la causa.

2. En sede legislativa:

2.1. Valorar la posibilidad de establecer un modelo estrictamente taxativo para los casos de VPMG.

2.2. Considerar la incorporación de otros medios de prueba, además de la documental y la técnica, en los procedimientos vinculados a VPMG.

V. FUENTES DE CONSULTA

Anderson, T. *et. al.* (2015) *Análisis de la prueba.* Traducción coordinada por Flavia Carbonell y Claudio Agüero, Marcial Pons.

Ansolabehere, K. *et. al.* (2017). Manual de análisis de contexto para casos de violaciones a los derechos humanos. FLACSO México, *International Bar Association's Human Rights Institute.*

Araya, M. (2020). Género y verdad. Valoración racional de la prueba en delitos de violencia patriarcal. Revista de estudios de la justicia, núm. 32. Género y verdad: Valoración racional de la prueba en los delitos de violencia patriarcal - Dialnet (unirioja.es).

Arena, J. (2022). Estereotipos normativos y autonomía personal. *Manual sobre los efectos de los estereotipos en la impartición de justicia.* Arena, J. (coord.). Dirección General de Derechos Humanos de la Suprema Corte de Justicia de la Nación.

Atienza, M. (2013). *Curso de argumentación jurídica*, Trotta.

Barrère, M. Á. (2008). *Iusfeminismo y derecho antidiscriminatorio: Hacia la igualdad por la discriminación*. En R. M. Mestre i Mestre (Coord.), Mujeres, derechos y ciudadanías (pp. 45–72). Tirant lo Blanch.

Bourdieu, P. (1996). Violencia simbólica. *Revista Latina de Sociología*, *2*(1), 1-4. *https://doi.org/10.17979/relaso.2012.2.1.1203*.

Cook, R. y Cusack, S. (2009). *Gender Stereotyping: Transnational Legal Perspectives*, *University of Pennsylvania Press*. Citado por la traducción al español de Andrea Parra, *Estereotipos de género. Perspectivas Legales Transnacionales*, Profamilia.

De Paula, V. *La carga de la prueba en el proceso civil. De la carga al deber de probar.* Traducción al español de: Laura Criado Sánchez, Marcial Pons.

De Paula, V. (2023). *La prueba documental. Del documento a los documentos y del soporte a la información*, Traducción al español de: Laura Criado Sánchez, Marcial Pons.

Ferrajoli, L. (2010). El principio de igualdad y la diferencia de género. En J. A. Cruz Parcero & R. Vázquez (Coords.), Debates constitucionales sobre derechos humanos de las mujeres (pp. 31–58). Suprema Corte de Justicia de la Nación; Fontamara.

Ferrer, J. (2007). *La valoración racional de la prueba*, Marcial Pons.

Ferrer. J. (2019). La carga dinámica de la prueba, Entre la confusión y lo innecesario, en: Contra la carga de la prueba, Nieva, J. *et. al.*, Marcial Pons.

Ferrer. J. (2021). *Prueba sin convicción. Estándares de prueba y debido proceso.* Marcial Pons.

Fuentes, O. (2022). Perspectiva de género y enjuiciamiento, en: La prueba de la violencia de género y su problemática judicial, Cerrato. E. (Dir.), La Ley.

Gardiner, G. (2018). *Evidentialism and Moral Encroachment*, *Oxford University*.

Gascón, M. (2005). *Los hechos en el derecho*. Marcial Pons.

González, D. (2005). *Quaestio Facti. Ensayos sobre prueba, causalidad y acción.* Palestra.

González, D. (2022). Filosofía de la mente y prueba de los estados mentales: Una defensa de los criterios de «sentido común», *Quaestio Facti*, *Revista Internacional sobre Razonamiento Probatorio*, núm, 3. *https://www.quaestiofacti.com/es/paginas/numeros*.

Mahoney, K. (1997). Enfoques canadienses a la igualdad de derechos y a la equidad de género en los estrados judiciales, en: Derechos humanos de la Mujer, Profamilia.

Pou, F. (2014). Argumentación judicial y perspectiva de género, en: Interpretación y argumentación jurídica en México, Cruz, J., *et. al.*, (coord.), Fontamara.

Ramírez, J. (2020). El testimonio único de la víctima en el proceso penal desde la perspectiva de género. *Quaestio Facti*, *Revista Internacional sobre Razonamiento Probatorio*, núm, 1. *https://www.quaestiofacti.com/es/paginas/numeros*.

Saba, R. (2007). «(Des)igualdad Estructural», en Roberto Gargarella y Marcelo Alegre (coordinadores), *El derecho a la igualdad. Aportes para un constitucionalismo igualitario*, Lexis Nexis.

Serrano, S. y Ortega A. (2016). La igualdad y la no discriminación. *Conceptos básicos y método, Guía de estudio de la materia principio de igualdad y no discriminación y la perspectiva de género*, Facultad Latinoamericana de Ciencias Sociales.

Serrano, S. (2019). Argumentar con PEG. Hechos y derechos: la recaracterización. *Guía de estudio de la materia argumentación desde la perspectiva de género*, Facultad Latinoamericana de Ciencias Sociales.

Taruffo, M. (2020). *La prueba de los hechos.* Traducción al español de Jordi Ferrer Beltrán, Marcial Pons.

Taruffo, M. (2010). *Simplemente la verdad.* Traducción al español por Daniela Accatino Scagliotti, Marcial Pons.

Vázquez, C. (2019). Técnica legislativa del feminicidio y sus problemas probatorios. *Doxa. Cuadernos de Filosofía del Derecho*, núm. 42. *https://doi.org/10.14198/DOXA2019.42.09*.

Vázquez, C. y Fernández, M. (2022). La conformación del conjunto de elementos de juicio: admisión de pruebas. *Manual de Razonamiento Probatorio*. Ferrer, J. (coord.). Dirección General de Derechos Humanos de la Suprema Corte de Justicia de la Nación.

Legislación

Legislación nacional

México. Constitución Política de los Estados Unidos Mexicanos, 5 de febrero de 1917, última reforma 15 de abril de 2025. *https://www.scjn.gob.mx/sites/default/files/cpeum/documento/cpeum.pdf*.

México. Ley General de Acceso de las Mujeres a una Vida Libre de Violencia, 1 de febrero de 2007, última reforma 16 de diciembre de 2024. *Suprema Corte de Justicia de la Nación*

México. Ley General de Instituciones y Procedimientos Electorales, 23 de mayo de 2014, última reforma 14 de octubre de 2024. *Suprema Corte de Justicia de la Nación*

Legislación internacional y extranjera

Bolivia. Ley Contra el Acoso y la Violencia Política Hacia las Mujeres, 28 de mayo de 2012. *https://www.ilo.org/dyn/natlex/docs/ELECTRONIC/90299/104007/F226460565/BOL90299.pdf*.

Organización de los Estados Americanos. Ley Modelo Interamericana para Prevenir, Sancionar y Erradicar la Violencia contra las mujeres en la vida política, octubre de 2016. *https://www.oas.org/es/mesecvi/docs/LeyModeloViolenciaPolitica-ES.pdf.*

Jurisprudencia y Tesis Aisladas

Suprema Corte de Justicia de la Nación

Jurisprudencia de la Primera Sala (México) 1a./J. 22/2016 (10a.), «Acceso a la justicia en condiciones de igualdad. Elementos para juzgar con perspectiva de género», Décima Época, *Gaceta del Semanario Judicial de la Federación*, Libro 29, abril de 2016, Tomo II, p. 836 (registro digital: 2011430).

Tribunal Electoral del Poder Judicial de la Federación

Tesis del Tribunal Electoral del Poder Judicial de la Federación (México) XLV/2002, «Derecho administrativo sancionador electoral. Le son aplicables los principios del ius puniendi desarrollados por el derecho penal», Tercera Época, Justicia Electoral. Revista del Tribunal Electoral del Poder Judicial de la Federación, Suplemento 6, Año 2003, pp. 121–122.

Jurisprudencia del Tribunal Electoral del Poder Judicial de la Federación (México) 18/2023, «Reversión de la carga probatoria. Procede en casos de violencia política en razón de género a favor de la víctima ante la constatación de dificultades probatorias» (Sala Superior, Séptima Época), aprobada en sesión pública de 24 de mayo de 2023. Gaceta de Jurisprudencia y Tesis en materia electoral, Año 16, núm. 28, número especial 18, 2023, pp. 33–35.

Jurisprudencia del Tribunal Electoral del Poder Judicial de la Federación (México) 22/2024, «Estereotipos de género en el lenguaje. Metodología para su análisis» (Sala Superior, Séptima Época), aprobada en sesión pública de 29 de mayo de 2024. Gaceta de Jurisprudencia y Tesis en materia electoral, Año 17, núm. 29, 2024, pp. 101–103.

Sentencias

Suprema Corte de Justicia de la Nación

Amparo directo en revisión 3186/2016, resuelto por la Primera Sala de la Suprema Corte de Justicia de la Nación el 1 de marzo de 2017.

Sala Superior del Tribunal Electoral del Poder Judicial de la Federación

SUP-JDC-12624/2011; SUP-JDC-383/2017; SUP-REP-252/2018; SUP-REP-250/2018; SUP-REC-91/2020 y acumulado; SUP-REC-133/2020; SUP-REP-245/2020; SUP-REP-394/2021;

SUP-REP-477/2021; SUP-REC-1861/2021; SUP-REC-2214/2021; SUP-JE-122/2022; SUP-REP-160/2022; SUP-REC-200/2022; SUP-REP-245/2022 y acumulados; SUP-REP-401/2024 y acumulados.

Sala Regional Especializada del Tribunal Electoral del Poder Judicial de la Federación

SRE-PSC-2/2021; SRE-PSC-102/2021; SRE-PSC-157/2021; SRE-PSC-195/2021; SRE-PSC-17/2022; SRE-PSC-41/2022; SRE-PSC-45/2022; SRE-PSC-94/2022; SRE-PSC-118/2022; SRE-PSC-5/2023; SRE-PSC-94/2024.

Conferencias

Conferencia impartida por Jordi Ferrer Beltrán el 26 de noviembre de 2019, organizada por la Dirección General de Derechos Humanos de la Suprema Corte de Justicia de la Nación, disponible en: *https://www.youtube.com/watch?v=SlsoI3WQLy4*.

CONJETURAS Y REFUTACIONES

Quaestio facti. Revista Internacional sobre Razonamiento Probatorio
Quaestio facti. International Journal on Evidential Legal Reasoning
Sección: Conjeturas y refutaciones
2026 I 10 pp. 363-384
Madrid, 2026
DOI: 10.33115/udg_bib/qf.i10.23299
Marcial Pons Ediciones Jurídicas y Sociales
© Anne Ruth Mackor
ISSN: 2604-6202
Recibido: 14/11/2025 | Aceptado: 16/12/2025 | Publicado *online*: 28/01/2026
Editado bajo licencia Reconocimiento 4.0 Internacional de Creative Commons

BAYESIAN MODELLING OF CRIMINAL CASES AS A WHOLE A PHILOSOPHICAL REFLECTION ON DUTCH CASE LAW

Anne Ruth Mackor*

ABSTRACT: The Bayesian model is gaining ground as a rational method for evaluating evidence in criminal law. The central question of this article is whether the Bayesian model can provide guidance for the judicial evaluation of evidence. More specifically, it focuses on the use of Bayesian modelling of complete criminal cases. Section 2 explicates the model, discussing an example of how experts from the Netherlands Forensic Institute (NFI) apply it in their reports. Section 3 addresses the question of whether non-forensic experts and judges can or even should use the Bayesian model as well. The focus of the rest of this article is on Bayesian modelling of criminal cases as a whole. In section 4, I discuss Dutch case law on the use of such integral modelling. The Netherlands appears to be internationally unique in that this question has been addressed in case law. Since the Bayesian model is worldwide gaining ground in forensic sciences and in theorizing about criminal evidence law, a discussion of Dutch case law can be relevant for legal scholars and legal philosophers from other, in particular from comparable continental legal systems. Section 5 offers a further legal-philosophical and epistemological reflection on the nature and functions of Bayesian modelling of criminal cases as a whole. The contribution concludes with recommendations for further research. Part of this further research is being carried out within the research project Preventing Miscarriages of Justice.

KEYWORDS: criminal law; case law; bayesian model; bayesian networks

* Anne Ruth Mackor is professor of philosophy of law at the Faculty of Law of the University of Groningen. a.r.mackor@rug.nl She is the PI of the NWO research project Prevention Miscarriages of Justice (PMJ) (no. 406.21.RB.004). This contribution was made possible with financial support from this project. See *https://preventingmiscarriagesofjustice.wordpress.com/*

I thank the members of the PMJ group and Henry Prakken for their helpful comments.

SUMMARY: 1. INTRODUCTION: EVIDENCE IN CRIMINAL LAW.— 2. THE BAYESIAN MODEL.— 3. THE SCOPE OF THE BAYESIAN MODEL IN CRIMINAL CASES.— 4. DUTCH CASE LAW ON BAYESIAN MODELLING OF CRIMINAL CASES AS A WHOLE: 4.1. The case of 'The six of Breda'; 4.2. Four later court rulings; 4.3. The silence of the Dutch Supreme Court.— 5. FURTHER CONSIDERATIONS REGARDING BAYESIAN MODELLING OF CRIMINAL CASES AS A WHOLE.— 6. CONCLUSION AND RECOMMENDATIONS.— REFERENCES

1. INTRODUCTION: EVIDENCE IN CRIMINAL LAW

Article 338 of the Dutch Code of Criminal Procedure requires that the judge be *convinced* by the content of the lawful evidence presented during the trial that the defendant committed the offence with which he or she is charged, before the judge may proceed to convict the defendant. Because of the term lawful evidence, Dutch criminal law is referred to as a *legal* system of proof. Because of the fact that a judge is not obliged to convict if there is sufficient lawful evidence, it is referred to as a *negative* system of proof. Accordingly it is a negative legal system of proof. However, as the law of evidence does not give further guidance to the judge's conviction, it is usually considered a *free* system of proof. As a consequence, Dutch criminal law does not so much leave the judge free, but rather helpless. It is unclear how the judge can or should rationally evaluate evidence, i.e. how and to what extent the content of the various lawful means of evidence can or should contribute to his conviction.

The Bayesian model is gaining ground in forensic sciences and in theorizing about criminal evidence law as a rational method of evaluating evidence. This article focuses on the question of whether the Bayesian model can provide guidance for the judicial evaluation of evidence. More specifically, it focuses on the use of Bayesian modelling of complete criminal cases. Section 2 explains the model, discussing an example of how experts from the Netherlands Forensic Institute (NFI) apply it in their reports. Section 3 addresses the question of whether non-forensic experts and judges can or even should use the Bayesian model. We will distinguish between the following questions. Can forensic and/or non-forensic experts use the Bayesian to determine the probative force of individual evidence? Can the Bayesian model be used for an integral modelling of a complete case as a whole? If so, should this be done by an expert or by the court itself?

The focus of the rest of this article is on the latter question, i.e. Bayesian modelling of criminal cases as a whole. In section 4, I discuss Dutch case law on the use of such integral Bayesian modelling. The Netherlands appears to be ahead of other countries, insofar as this question has been addressed in Dutch case law. Because the Bayesian model is gaining ground as a rational method for evidence evaluation, a discussion of Dutch case law can also be relevant for legal scholars and legal philosophers from other, in particular from comparable continental, legal systems. Section 5 offers a legal-philosophical reflection on the nature and functions of Bayesian mod-

elling of criminal cases as a whole. The article concludes with recommendations for further research. Part of this further research is being carried out within the research project Preventing Miscarriages of Justice. [1]

2. THE BAYESIAN MODEL

Since the 1970s, the Bayesian model has been gaining ground as a rational method of evaluating evidence in forensic sciences and in theorizing about criminal evidence law. [2] For the past 20 years, it has become widely accepted among European forensic scientists as a method for rationally evaluating evidence. Let me start with a simple example. It is known from statistics that it rains more than half of the days of the year in the Netherlands. [3] However, after looking outside and seeing bright sunshine ('evidence' 1) and hearing the weather forecast predicting "dry everywhere" ('evidence' 2), I will lower my estimate of the probability of rain for today. This is something we already do spontaneously and intuitively; the Bayesian model provides a precise method for this.

Put simply, the model prescribes that we first estimate the prior odds, that is, prior to the evidence, that it will or will not rain. In the example we can do this on the basis of official statistics. [4] When these are not available, we will base our estimate on our own experience or informally on 'facts of general knowledge'. The next step is to estimate which hypothesis (in the example: rain or no rain) makes the evidence (bright sunshine plus weather forecast) more likely. This second estimate is called the likelihood ratio (LR). The greater the LR, the stronger the evidence. Other words for likelihood ratio are therefore "evidential strength", "probative strenght", "evidential force" and "probative force". The Bayesian model offers a method to derive the posterior odds that it will or will not rain today from the prior odds (in this example: based on the statistics on the number of days of rain per year) and the LR of the evidence (in this example: bright sunshine plus weather forecast). In a 'formula': prior odds x likelihood ratio = posterior odds. [5]

---

[1] See note *.

[2] Note that the Bayesian model is not the only method for rational evaluation of evidence. H. Prakken, F.J. Bex & A.R. Mackor Editors' Review and Introduction: Models of Rational Proof in Criminal Law*, Topics in Cognitive Science* 12 (2020) 1053–1067 offer an overview of three main approaches, viz. argumentation-based approaches; explanation-based, more in particular story-based approaches; and probabilistic, in particular Bayesian, approaches.

[3] *https://www.knmi.nl/over-het-knmi/nieuws/het-regent-niet-vaker-wel-harder*

[4] An important question is which statistics this assessment should be based on. In the example: on the statistics for the entire year, for this month, or for this specific day? I will leave this problem, which is known as the reference class problem, aside in this paper.

[5] NFI *Vakbijlage Waarschijnlijkheidstermen* [*Technical Supplement Probability Terms*], 2017. *https://www.forensischinstituut.nl/publicaties/publicaties/2017/10/18/vakbijlage-waarschijnlijkheidstermen*, p 3.

In a similar way the Bayesian model can be applied in criminal cases. The focus is usually on the question of the posterior probability that the defendant committed the alleged offence in light of the prior probability and the probative force of all the evidence (the LR's).[6] Since around 2005, the Netherlands Forensic Institute (NFI) writes its reports in accordance with the Bayesian model.[7] In 2017, the NFI published a technical supplement explaining the model and its application in NFI reports.[8] NFI researchers report on the LR, that is, on the probability of finding the research results in the light of at least two hypotheses. Consider the following example.[9]

> A researcher compares a shoe print with the suspect's shoe. He formulates the following hypotheses.
> H1 The suspect's shoe caused the print.
> H2 Another shoe with a similar profile and size caused the print.
> After forensic investigation, the conclusion is:
> The results of the investigation are much more likely if the suspect's shoe caused the mark than if another shoe with a similar profile and size caused the mark.

The report gives the LR of the shoe mark. In this example, the LR is "the results of the investigation are much more likely if H1 is true than if H2 is true". "Much more likely" means that the results are 100 to 10,000 times more likely.[10] In other words, the probability of finding the shoe mark with this profile and size is 100 to 10,000 times greater if the mark was made by the suspect's shoe than if it was made by any other shoe. In this example, the LR is qualitative, i.e. expressed in words. A quantitative LR, expressed as a specific number, is possible if relevant data is available from representative samples or experiments.

Next, it is up to the court to combine the LR from the NFI report with all other information. This is a complex task. First, the court must estimate the prior probabilities of both hypotheses. In other words, the court must estimate the probability, prior to the evidence, that the suspect's shoe and not another shoe made the mark. Sometimes that estimation can be based on statistics. We saw that this was possible when estimating the prior probability of rain. However, in criminal cases reliable statistics are often lacking, and then the estimation is necessarily subjective. That is also why experts do not offer estimates of the prior probability and leave that task to

---

[6] Another important probability assessment in criminal law is that of the risk of recidivism.

[7] H. Prakken & R. Meester, Bayesiaanse analyses van complexe strafzaken door deskundigen. Betrouwbaar en zo ja: nuttig? [Bayesian analyses of complex criminal cases by experts. Reliable and, if so, useful?] *Expertise en Recht* 2017-5, 185-197.

[8] NFI 2017.

[9] NFI 2017, p. 2.

[10] NFI 2017, p. 6, table 2. When it comes to hypotheses about the source of DNA traces, the results are often "much more likely" (10,000 to 1,000,000 times) or even 'extremely much more likely' (more than 1,000,000 times) under the hypothesis that the trace comes from the suspect than under the hypothesis that it comes from any other person. However, much evidence has a much smaller LR. For example, 'slightly more likely' means that finding the evidence is 2 to 10 times more likely, and 'more likely' means 10 to 100 times more likely.

the court. After all, in such cases they do not have the specific knowledge or expertise on which to base such an estimation.[11] Secondly, the court must estimate the probative force of all other evidence. The NFI can provide a LR for some of the evidence, but there is also a lot of evidence for which scientific knowledge is lacking and on which experts are therefore unable to report. For that evidence, the court will have to estimate the probative force by itself.[12] Finally, the court must combine the NFI's LR with the probative force of all other evidence, such as witness statements, DNA, glass, fibre or gunpowder traces and the relevant priors.

It is for the latter task, the combination of priors and LR's, that the Bayesian model offers a precise mathematical method. The method can also help to prevent reasoning errors, such as the notorious prosecutor's fallacy.[13] Again, take the example of the shoe print. The conclusion of the above-mentioned NFI-report was: "The results of the investigation are much more likely if the suspect's shoe caused the mark than if another shoe with a similar profile and size caused the mark."

However, many people, including judges, read the conclusion as: "The probability that the shoe print came from the suspect's shoe (the hypothesis), given the match (the evidence), is much greater than the probability that the print was made by another shoe (the alternative hypothesis)." This interpretation is incorrect, however. The mistake is that hypothesis and evidence are reversed. In general terms: the NFI reports on the LR, which is the conditional probability of finding the evidence if the hypothesis is true. However, many people, including judges, read it the other way around as the posterior, the conditional probability that the hypothesis is true, given the evidence. A classic example can illustrate that this can be a major mistake: the

[11] The Amsterdam Court of Appeal of 12 April 2021 (ECLI:NL:GHAMS:2021:986), quotes a forensic expert who explains that: "... the NFI deliberately refrains from making a priori assessments, because the NFI cannot determine those probabilities. The NFI never has access to all the information in a case, because a case always contains information that falls outside the NFI's area of expertise. The court must determine what the a priori probability is, given the established facts before and during the incident. The NFI's statements only relate to the discovery of ... findings as to whether a hypothesis is correct, compared to an alternative hypothesis. In view of the above, the Bayesian method of reasoning is pure, because this method allows experts to remain within their own area of expertise."

[12] For example, a judge might ask how much more likely it is that the defendant filled up with a small amount of petrol shortly before a certain fire took place (evidence) if the defendant is the perpetrator (H1) than if he is not the perpetrator (H2). If no statistics are available, judges would have to rely on their own experience or on so-called 'facts of general knowledge'. This example is taken from the conclusion of Advocate General P.M. Frielink of 25 May 2025 (ECLI:NL:PHR:2025:378).

[13] The term is taken from William C. Thompson and Edward L. Shumann, "Interpretation of Statistical Evidence in Criminal Trials: The Prosecutor's Fallacy and the Defense Attorney's Fallacy." *Law and Human Behavior* 2 (3): 167–87. Also see C. Dahlman, A systematic account of probabilistic fallacies in legal fact-finding, *The International Journal of Evidence & Proof* 2024/1, pp. 45-64 and Hylke Jellema and Christian Dahlman, Probabilistische drogredenen in de rechtszaal: een systematisch overzicht [Probabilistic fallacies in the courtroom: a systematic overview], *Expertise en Recht* 2025(3), 64-74 for an overview and explanation of the most important fallacies.

conditional probability that an animal has four legs if it is a cow is very high, whereas the conditional probability that an animal is a cow if it has four legs is very low.

Because of the complexity of the details of the Bayesian model and the scope of this article, the question of exactly how the Bayesian model should be applied in a mathematical manner cannot be addressed. However, the above explanation should be sufficient, also for legal scholars without knowledge of the Bayesian model, to understand the discussion in the rest of this article.[14]

3. THE SCOPE OF THE BAYESIAN MODEL IN CRIMINAL CASES

In the previous section, I gave an example of the application of the Bayesian model to forensic evidence. Because the Bayesian model can not only be used quantitatively, that is, in numbers (e.g. an LR of 950), but also qualitatively, that is, verbally (e.g. "much more likely"), it is in principle suitable for application to all evidence, including evidence for which no relevant data from representative samples or experiments are available.

It has been argued that legal psychologists should also report according to the Bayesian model.[15] For example, a psychologist could report that it is 'slightly more likely' or 'more likely' that this witness recognised the suspect as the perpetrator (H1) than that he is not the perpetrator (H2). Again, note that the psychological expert also reports on the LR, the conditional probability of the evidence given the hypotheses, not on the posterior, the conditional probability of the hypotheses given the evidence.

If, in principle, a qualitative or quantitative LR can be assigned to all evidence, the question arises as to whether judges themselves may or even should evaluate evidence according to the Bayesian model, particularly evidence that experts cannot report on because there is insufficient information available for a scientifically sound estimation. If the answer to these questions is affirmative, then a logical follow-up question is whether the Bayesian model can be applied to analyse a criminal case as a whole.

This follow-up question is the subject of debate. A first question is what is meant by a case as a whole. 'As a whole' does not mean that all the information in the file is modelled, but it does mean that at least the central or ultimate (alternative) hypotheses (in short: minimally the guilt and innocence hypotheses) and the most important evidence are analysed in relation to each other. An important question is who should

---

[14] The Bayesian method is briefly explained in NFI 2017, in particular pp. 3-5 (in Dutch).

[15] See Rassin, Arbiyah, Boskovic., Otgaar., & Merckelbach, (2022). Likelihood ratios in psychological expert opinion, and their reception by professional judges. *The International Journal of Evidence & Proof*, *26*(4), 325-341. *https://doi.org/10.1177/13657127221119545*. For an example of the use of LR in a forensic psychological report, see the conclusion of A-G P.C. Vegter of 16 June 2020 (ECLI:NL:PHR:2020:560). More on this in section 4.3.

or may perform such an integral Bayesian analysis.[16] This presents a dilemma.[17] According to Dutch law it is the task of the court to select and evaluate the evidence, but judges are not experts in Bayesian modelling.[18] It is therefore, at least at present, unlikely that judges are able to carry out a Bayesian analysis on their own. An obvious alternative is for an expert to carry out such an analysis for the court.

However, there are at least two arguments against this. First, the expert would not only have to be an expert in Bayesian modelling, but also a 'material' expert in all types of evidence for which experts have not provided LR. Of course, no such 'omniscient' experts exist. A second objection is that even if we could find such 'omniscient' experts, they would be sitting in the judge's chair as a result of the choices they make in selecting, evaluating and combining the hypotheses and evidence. This would be unacceptable from a rule of law perspective.

An alternative is for judges themselves to select and evaluate hypotheses and evidence, but to be assisted in their Bayesian modelling, for example by the court's forensic advisor.[19] However, this does not solve the problem, at least not entirely. After all, even then there is a risk that the forensic adviser's modelling proposals will exert undue influence on the judge's evidential decision.

All this brings us to the following four questions:

1. Can the Bayesian model be used to determine the probative force (the LR) of individual evidence?

a. By forensic experts?

b. By non-forensic experts, for example by legal psychologists?

---

[16] I take the term "integral" from the judgment of the Court of Appeal in The Hague of 14 October 2015 (ECLI:NL:GHDHA:2015:2860) that will be discussed in section 4.1.

[17] More on this dilemma in A.R. Mackor & R. Schutgens, De kansendeskundige als lekenrechter: Artikel 116 Grondwet, 2022, [The probability expert as lay judge: Article 116 of the Constitution], in B. Rijpkema & A. Ellian (eds.). *Een nieuw commentaar op de grondwet*. Boom Publishers Amsterdam, 2022, 451-466. See also A.R. Mackor, Risks of Incorrect Use of Probabilities in Court and What to Do about Them, in: *Risk and Responsibility in Context*, in A. Placani & S. Broadhead (eds.). New York and London: Routledge, 2024, pp. 94-108.

[18] In fact, research shows that judges regularly misunderstand Bayesian reports. See Jan de Keijser & Henk Elffers, 'Understanding of Forensic Expert Reports by Judges, Defense Lawyers and Forensic Professionals', *Psychology, Crime & Law* 2012, 18 (2): 191-207 for an experiment with, among others, Dutch judges. For examples of the prosecutor's fallacy in Dutch case law, see, for example, H. Prakken, Kansoordelen door deskundigen: over "logisch" rapporteren en wat daarbij mis kan gaan. [Probability assessments by experts: on "logical" reporting and what can go wrong.] *Ars Aequi* 67 (2018), 740-747; see also the NFI professional supplement pp. 7-8.

[19] Forensic advisors are appointed at all Dutch criminal courts. For more information about the role of forensic advisor, see Rosanne de Roo, Wanda Remijn, Jaimy Meeuwissen, Jeltje Kruithof-van Esch, Suzanne van der Heijden, Michelle Claushuis & Laura van Blijswijk-Kieftenbeld, Forensisch adviseurs actief bij alle gerechten in Nederland: tijd voor een terugblik [Forensic advisors active in all courts in the Netherlands: time for a review], *Expertise en Recht* 2021-6, 234-239.

2. Can the Bayesian model be used for an integral modelling of a case as a whole?

a. By an expert?
b. By the court itself, possibly with the help of the court's forensic advisor?

In the following section, we discuss the response of Dutch case law to question 2a. Section 5 discusses question 2b from a broader philosophical perspective.

4. DUTCH CASE LAW ON BAYESIAN MODELLING OF CRIMINAL CASES AS A WHOLE

4.1. The case of 'The six of Breda'

At first glance, these legal-philosophical, constitutional and epistemological questions may seem purely theoretical. However, they have already been addressed several times in Dutch case law, viz in cases in which the Public Prosecution or the defence had submitted a report containing such an integral analysis. According to the Court of Appeal The Hague, this happened for the first time in its ruling of 14 October 2015 (ECLI:NL:GHDHA:2015:2860) in the review case of the six from Breda.[20] The ruling explicitly addresses the use of a report submitted by the Public Prosecution containing a integral Bayesian analysis of the case.[21] The Court did not take the conclusion(s) of the report into consideration in its assessment of the evidence. It justified its decision as follows. The Court states that it has

> " […] found that there is insufficient scientific consensus on the question of whether the method of (classical) Bayesian analysis used by [expert 1] can be reliably applied to and in complex criminal cases such as the present one, in which there are many findings that cannot always be regarded as independent of each other. […]"
> "Moreover, there appear to be considerable differences of opinion among experts as to which (and how) findings in complex criminal cases should be included in the analysis, how the (a priori) probabilities of the findings can be (reliably) determined or estimated, and what expertise is necessary for that determination or estimation. This leads the court to conclude that, given the current state of scientific knowledge, the reliability of the aforementioned method and, therefore, of the results obtained from its use, is too uncertain for actual application in a complex criminal case such as the one at hand."

Which of the four questions that were distinguished in section 3 does the Court answer in this ruling? The wording 'application of Bayesian analysis in complex criminal cases' is somewhat misleading. This case does not concern the application of

[20] Under 2.2, the court of appeal writes that "... to the court's knowledge, such a more comprehensive Bayesian analysis has not yet been used in a criminal case."

[21] The heading of 2.2 reads: 'The admissibility and reliability as evidence of the Bayesian report', whereas this concerns not evidence but an expert report on the evidence. This wording may be a consequence of the defence's position, cited in 2.2, that the report 'is completely unusable as either a guideline or as evidence in the assessment of the present criminal case'.

the Bayesian model by a forensic expert to individual evidence in a complex case (question 1a), but an integral Bayesian analysis of a complex case as a whole by an 'omniscient' expert (question 2a).

The Court of Appeal rightly notes that in (an integral analysis of) a complex case many choices have to be made and that this requires many different types of expertise. In this case, the expert must estimate the prior probabilities of the hypotheses and the LR's of all the evidence. The court of appeal seems to incorrectly conclude from this correct insight that the reliability of the Bayesian method as such is too uncertain for application.

However, as was already stated in section 2, the method itself is not disputed among forensic experts. The reliability of the Bayesian method comes into question in the selection and evaluation of the hypotheses, the evidence and their interdependencies; not in the method as such. The reliability of the application is particularly questionable when an expert performs an integral analysis (question 2a), because there is a high probability that his selection and evaluation will not, or at least not entirely, correspond to that of the court, thereby placing the expert on the court's chair.[22] [23]

However, this argument against an integral analysis does not hold if judges themselves apply the Bayesian method, for example during deliberations in the court chamber (question 2b). The argument is even less sound if judges do not rely solely on this method, but primarily use it as a means of inspection after they have arrived at a preliminary assessment of the evidence in a 'traditional' manner. They can compare their preliminary assessment of the evidence with the outcome of a more or less detailed integral Bayesian analysis.[24] If there are (significant) differences, this may

[22] More on this topic in section 3.

[23] In other words, if the inter-rater reliability of the application of the method is low, the expert takes the place of the judge. See also Prakken and Meester 2017. "Reliability means ... whether different people applying this method to the same case would arrive at the same outcome. The question is to what extent Bayesian network models of complex criminal cases can be reliable in this sense. ... In Bayesian modelling, at least the following choices must be made: 1. Formulating the hypotheses, 2. Selecting potentially relevant findings, 3. Determining statistical dependency relationships between the various findings given the hypotheses, 4. Estimating or positing the relevant probabilities, 5. Determining whether further potential findings could be relevant."

[24] In the Netherlands, an over-simplification of an integral Bayesian analysis has been informally referred to as "Bayes on a beer mat". Marjan Sjerps, forensic statistician at the Dutch Forensic Institute NFI, writes about this in a personal capacity: 'I think it can be useful for legal decision-makers ... to take a very rough look at what the most important pieces of evidence are, what the whole thing is worth approximately and how likely it is that the Public Prosecution Service is right. For example, for criminal judges who—after the usual deliberations—make this calculation, it can be an extra check to see whether the decision is reasonable." (personal e-mail correspondence, 13 August 2025). Also note that in such cases, support from a forensic advisor also seems less problematic, because there is already a draft evidence assessment that has been established in the 'traditional' manner and the question of whether and why the Bayesian modelling corresponds to or deviates from the 'traditional' evidence assessment is explicitly addressed.

give rise to critical reflection on possible errors of reasoning in the 'traditional' and/or Bayesian analysis. In this way, a Bayesian analysis can contribute to the 'sharpening function' of the reasoning in the evidential decision. In section 5 this option will be discussed more extensively.

4.2. Four later court rulings[25]

In four later court rulings, other courts have also decided on the use of the Bayesian method by an expert in the analysis of a case as a whole. On 23 May 2016, the Zeeland West-Brabant District Court (ECLI:NL:RBZWB:2016:3060), like the Court of Appeal The Hague, did not take the expert's report into consideration.

The district court writes:

> "Derksen's [the expert, ARM] calculation, which would appear to show that the defendant is, in short, probably innocent, was made according to Bayes' theorem. In the opinion of the court, this means that a controversial rule from probability theory has been used for the purpose of establishing the truth in criminal proceedings, the outcome of which depends to a large extent on the selection and evaluation of the evidence made by Derksen. The court does not rule out that this method, if applied correctly, can be an aid to finding the truth, but the ultimate answer to the question of whether a charged offence can be proven legally and convincingly is not a question that lends itself to probability theory. The answer to that question is and remains reserved for the judge. In addition, in this case, the court makes a different selection and assessment of the evidence than Derksen."

Like the court of appeal of the Hague, the district court answers question 2a in the negative. Unfortunately, this court too fails to make a clear distinction between the application of the Bayesian method to individual evidence (question 1) and to a criminal case as a whole (question 2). Also, when the district court writes that Bayes' theorem is not undisputed, it confuses —like the court of appeal of The Hague—the validity of the method as such with the inter-rater reliability of its application by an expert in the specific case.[26]

The district court corrects itself by noting that the Bayesian method may be an aid to find the truth, but then states without further argument that the ultimate question of proving guilt does not lend itself to probability theory. The question

---

[25] The case law on the use of Bayesian modelling in criminal cases was selected by means of the search terms "Bayes" (17 hits), "Bayesian" (17 hits) and "Bayesian" (94 hits). I examined all court rulings and conclusions of Advocates General at the Supreme Court in the field of criminal law up to 1 October 2025 that have been published on the website of the judiciary (*https://uitspraken.rechtspraak.nl/*). The total number of rulings is not the sum of these numbers; some rulings contain more than one term. In criminal law, the term "Bayesian" first appears in the judgment of the Court of Appeal in 's-Hertogenbosch of 24 September 2008 (ECLI:NL:GHSHE:2008:BF2188). Incidentally, the first hit in case law is not in criminal law, but in tax law. It concerns the conclusion of Advocate General Niessen of 14 March 2008 (ECLI:NL:PHR:2008:AU0837). The vast majority of case law concern the application of the Bayesian model to individual forensic evidence.

[26] See note 23 on inter-rater reliability.

of how the judge can rationally reach a verdict of guilt without using some form of probability theory remains unanswered.[27] Finally, the court also fails to distinguish between rejecting an integral Bayesian analysis by an 'omniscient' expert (question 2a) an integral Bayesian analysis by the court itself, whereby the court retains control over the selection and evaluation of the evidence (question 2b).

On 5 February 2018, the Court of Appeal in 's-Hertogenbosch (ECLI:NL:GHSHE:2018:421), in the appeal in this case, also did not take Derksen's reports into consideration. The Court of Appeal argued at length that Derksen's selection of facts was arbitrary and that his argument was incomprehensible in some respects. However, the arguments put forward by the court against the use of the Bayesian method in general are less convincing. Under the heading "Research reports by Prof. Dr T. Derksen", the court writes:

> "... the Bayesian method, a quantitative research method that views legal evidence as thinking about probabilities, has been subject to much criticism in the literature. First of all, this method assumes that, given the hypotheses, all means of evidence are statistically independent of each other. In addition, it is doubtful whether the probabilities of the components of the input in the Bayesian method can be expressed in reliable numbers. The use of numbers creates an appearance of objectivity. The a priori probability of a component to be introduced can only be estimated subjectively, and this subjective probability cannot be tested objectively. Finally, the question is whether degrees of belief must comply with the axioms of probability theory."

I will limit myself to a few observations. First, this court too fails to distinguish between question 1a, question 2a and question 2b. Second, is noteworthy that the court raises critical questions about the Bayesian method, but at the same time allows a Bayesian analysis of individual forensic evidence. Third, the court incorrectly states that the Bayesian method assumes that evidence is independent.[28] In fact, the strength of the Bayesian method is that it can deal with both dependent and independent evidence. More on this in section 5. Fourth, the court incorrectly equates the quantitative application of the Bayesian method with the Bayesian model perse. We have seen that the Bayesian model can also be applied qualitatively. Furthermore, the court refers to "a subjective probability that cannot be tested objectively", without explaining what it means by subjective and objective. The term "subjective" suggests arbitrariness, but that is not necessarily true.[29] Finally, the court ignores the fact that the current judicial 'free evaluation of evidence' is less transparent than a Bayesian analysis.

The court also makes two interesting remarks. The first is about the reliability of the numbers and the false appearance of objectivity numbers can create. I will discuss this topic in section 5. Second, the court poses the question whether degrees of belief must

---

[27] At least two alternatives approaches could be used, viz an argumentative or an explanation-based or scenario approach. See note 2.

[28] This argument only applies to a simplified method also known as "linear Bayes".

[29] Subjective in the Bayesian sense means that probabilities are not properties of reality, but of our minds. However, these subjective Bayesian beliefs need not be subjective in the sense of irrational. They can be based on scientific information and they can be discussed and assessed intersubjectively.

comply with the axioms of (Bayesian) probability theory. The answer to this question is not necessarily affirmative; other approaches exist.[30] However, the court does not make it clear which approach it would consider as an alternative to the Bayesian model. Moreover, now that the court does allow the Bayesian model for individual evidence (question 1), it is unclear how, in the court's view, these different approaches could or should be combined in a Bayesian analysis of a case as a whole (question 2).

The next case also concerns a Bayesian analysis of a criminal case as a whole. On 30 May 2017, the Court of Appeal in 's-Hertogenbosch (ECLI:NL:GHSHE:2017:2300) rejected the Public Prosecution's request to appoint Alkemade as an expert.[31] The court limited itself to a brief statement of reasons in line with both previous rulings:

> "As argued by the Public Prosecution and the defence, and as is also apparent from the literature on the subject, Bayesian analysis in complex criminal cases is not without controversy. The reliability of the Bayesian analysis method is still uncertain according to the current state of science."

This ruling also fails to make it clear that the problem with Bayes does not lie in the use of the Bayesian model in complex criminal cases as such. After all, it is precisely in complex cases that NFI experts regularly provide Bayesian analyses of forensic evidence (question 1a). Again, the problem in this case lies solely with an integral Bayesian analysis by an 'omniscient' expert (question 2a).[32]

Three years later, on 12 November 2020, the same court of appeal in 's-Hertogenbosch (ECLI:NL:GHSHE:2020:3483) cleared up one of the misunderstandings created by the rulings discussed above. The court wrote that:

> "[...] there is consensus on the use of the Bayesian method in calculating the probative force of forensic evidence alone at forensic institutes in the Netherlands and abroad. The defence counsel's argument apparently refers to the use of the Bayesian method to analyse the entire case. However, that situation does not apply in this case, as only the forensic evidence has been examined. The court therefore, like the district court, sees no reason to doubt the Bayesian method used by the NFI and the findings of the investigation carried out on the basis of that method."[33]

---

[30] One of the best known is the Dempster-Shafer theory of belief functions.

[31] In another case in which the same expert, again at the request of the Public Prosecution, had drawn up a Bayesian report, the Court of Appeal in 's-Hertogenbosch, 22 November 2016, ECLI:NL:GHSHE:2016:5165, did not pass judgement on the use of the Bayesian model as such. It merely states: 'In view of the objections raised by Prof. H. Prakken in the counter-investigation, the court did not have the report by Dr F.J.M. Alkemade contribute to the court's ruling on the evidence decision.' On these cases, also see H. Prakken, Rechters en Raadsheren over Bayes [Judges and Justices on Bayes]. In R. Horselenberg, V. van Koppen, & J. de Keijser (eds.), *Bakens in de Rechtspsychologie: Liber Amoricum voor Peter van Koppen,* Boom Uitgevers, 2020, 287-300.

[32] The court of appeal also explicitly states however: "The defence's assertion that the Public Prosecution should not be allowed to introduce a report by Dr Alkemade into the proceedings is not supported by law." The Court of Appeal reiterates this view in its ruling of 9 October 2018 (ECLI:NL:GHSHE:2018:4181). Accordingly, the admissibility of a report with an integral Bayesian analysis is not under discussion.

[33] Other courts had previously ruled in a similar vein. On 25 February 2015 (ECLI:NL:RBLIM:2015:1588*)*, the Limburg District Court ruled in the first instance in the same case

Thus, the court explicitly states that the Bayesian method is generally accepted for the evaluation of forensic evidence (question 1a) and—or so I would add—also in complex criminal cases. The court also states that this method is only open to debate when it is used to analyse a case as a whole (question 2). However, the ruling leaves open the question of whether the Bayesian method is acceptable when non-forensic experts, such as legal psychologists, apply it to individual pieces of evidence (question 1b). Nor does the court rule on the question of whether there is scope for an integrated Bayesian analysis when this is carried out by the judge himself (question 2b).

4.3. The silence of the Dutch Supreme Court

Unfortunately, so far only this ruling by the Court of Appeal in 's-Hertogenbosch, which has removed some of the ambiguities, has been issued. Two other cases in which the defence submitted critical questions about the use of the Bayesian method to the Dutch Supreme Court were disposed of by the Supreme Court on the basis of Article 81(1) of the Judicial Organisation Act.[34] Also neither of the two Advocates General at the Supreme court provided clarity in their conclusions.

In the first case (Supreme Court, 25 August 2020, ECLI:NL:HR:2020:1317), the defence criticises the use of the Bayesian method by a legal psychologist (question 1b). In the conclusion of 16 June 2020 (PHR, ECLI:NL:PHR:2020:560), the Advocate General dismisses this criticism by claiming that an evaluation of the reliability of the method is factual in nature and does not fit within the cassation proceedings. This view may not be legally contestable, but it can be critized on more fundamental grounds. The question of whether a method has been applied reliably in a specific case may be a question that must be answered by the trial court, but answering the question of whether a method as such is suitable for the specific type of application does not seem to be reserved exclusively for the trial court. In any case, the Advocate

---

that "…the method of investigation used by the NFI, namely the Bayesian method, is generally accepted in science and is widely regarded as a scientifically sound method of investigation". On 16 February 2016 (ECLI:NL:RBGEL:2016:825), the Gelderland District Court stated: "The experts have stated that the Bayesian method was used in the "Breda Six" case cited by the defence. However, according to the experts, in the present case, a Bayesian evaluation of the scientific probative force of a reported DNA match was used. The evaluation of the probative force therefore only concerns the DNA evidence. The comparison with the "Breda Six" is therefore incorrect.' On 19 December 2017, the District Court of East Brabant stated (ECLI:NL:RBOBR:2017:6556): 'Bayesian statistics have been used for years, both nationally and internationally, to analyse forensic evidence. It is a logically correct way of forming an expert opinion, explicitly avoiding making statements outside the area of expertise. The use of this form of statistics in, for example, the field of forensic DNA research has long been considered standard practice and is generally uncontroversial."

[34] Article 81(1) of the Dutch Code on the Judicial Organisation reads: "If the Supreme Court finds that a complaint lodged cannot lead to cassation and does not require a response to legal questions in the interest of legal uniformity or legal development, it may limit itself to this finding when stating the grounds for its decision."

General could have taken the opportunity to clarify that the application of Bayes by legal psychologists (question 1b) does not differ in principle from its application by forensic investigators (question 1a) and therefore—in line with case law, more specifically the above-mentioned ruling of the Court of Appeal in 's-Hertogenbosch of 12 November 2020—is in principle unproblematic.

In the second case (Supreme Court, 15 June 2021, ECLI:NL:HR:2021:788), the defence criticises the use of the Bayesian model, but does not make a clear distinction between the generally accepted use of Bayes by forensic experts on individual evidence (question 1a) and the controversial use of the Bayesian model for an integrated analysis of a case as a whole (question 2).[35] Here too, the Advocate General in his conclusion of 20 April 2021 (PHR, ECLI:NL:PHR:2021:391), fails to take the opportunity to provide clarity on this distinction and on the use of the Bayesian model in criminal cases. The Advocate General could minimally have made it clear that the use of the Bayesian method by forensic experts to determine the probative force of individual forensic evidence is undisputed.

5. FURTHER CONSIDERATIONS REGARDING BAYESIAN MODELLING OF CRIMINAL CASES AS A WHOLE

Partly in response to the case law from 2015, 2016 and 2017 that was discussed in section 4, a brief but heated debate has taken place in Dutch literature on the question of whether there should be room for Bayesian analyses of cases as a whole by experts. Prakken and Meester have explicitly spoken out against it.[36] At least for now, they only see a role for the Bayesian method in the analysis of individual (forensic) evidence and for qualitative and global use to recognise and prevent probabilistic

[35] The case in question only concerns the application by forensic experts to individual evidence, not an integral analysis of a case as a whole. Another and more important complaint by the defence concerns the fact that the DNA traces were only examined at source level (the question of whether the DNA originated from the suspect) and not also at activity level (the question of how what activity caused the DNA trace). See section 5 on this distinction.

[36] Prakken and Meester 2017. In their article, the authors discuss whether Bayes' rule can be applied in a scientifically reliable and useful manner in court to analyse complex criminal cases as a whole and whether such use of Bayes' rule in court is practically possible and, if so, under what conditions. Their article has also been noted in practice; the conclusion of Advocate General Hofstee of 20 April 2021 (ECLI:NL:PHR:2021:391) shows that the defence referred to their article. Incidentally, Prakken was a counter-expert in two cases in which Alkemade acted as an expert, and in that capacity he sharply criticised the content of both of Alkemade's reports and pointed out probabilistic errors in reasoning therein. See Frans Alkemade, Makkers, staakt uw wild geraas. Een oproep tot de Nederlandse deskundigen in forensic reasoning om een consensus te formuleren [Friends, cease your wild clamour. A call to Dutch experts in forensic reasoning to formulate a consensus], *Expertise en Recht* 2018 (1), 26-35 for a response to Prakken and Meester 2017.

errors in reasoning.[37] However, Prakken and Meester do not address the question of who could perform an integral Bayesian analysis, at what stage, and for what purpose. In order to arrive at a balanced judgement on the use of Bayesian analyses of criminal cases as a whole, we must minimally distinguish between the following questions.[38]

1 Who performs the Bayesian analysis?

At a minimum, a distinction must be made between:

A Expert or forensic advisor

B Judge(s)

C Judge(s) with the support of a forensic advisor

2 At what stage is the analysis used?

At a minimum, a distinction must be made between:

A Hearing and proceedings[39]

B Court chamber

C Judicial ruling

3 What is the function of the analysis?

At a minimum, a distinction must be made between:

A Contributing to the investigation at the hearing

B Sharpening the judicial thought process and, related to this, internal check of the evidence decision during the court chamber's meeting

C Contributing to the justification of the evidence decision in the ruling[40]

As part of the research project Preventing Miscarriages of Justice[41], our research group is studying the added value of a combined use of the scenario approach and the Bayesian model in the evaluation of evidence in criminal cases.[42] Here I will limit myself to a discussion of the research of our group into the use of the Bayesian model. Dahlman has proposed a new systematisation of probabilistic thinking errors.[43] Based

---

[37] For examples of a qualitative and global analysis of several court rulings, see R. Meester and L. Stevens, Correct redeneren: wat een Bayesiaanse analyse zegt over vier recente strafrechtelijke uitspraken [Correct reasoning: what a Bayesian analysis says about four recent criminal court rulings]. *Expertise en Recht*, 2021(3), 112-121.

[38] These distinctions are not intended to be exhaustive. I limit myself to the function of Bayesian analysis for the judiciary. I do not discuss its role for investigation, prosecution and defence.

[39] According to Dutch case law, expert reports containing a Bayesian analysis of a case as a whole may be brought into the proceedings (2a). See note 32.

[40] Above we have seen that the Court of Appeal in 's-Hertogenbosch states that it did not have the report on the integral Bayesian modelling contribute to its ruling on the evidence decision. See note 31.

[41] See note *.

[42] On the scenario approach, see A.R. Mackor and P.J. van Koppen, The scenario theory about evidence in criminal law. In: Christian Dahlman, Alex Stein, and Giovanni Tuzet (eds.), *Philosophical Foundations of Evidence Law*, Oxford University Press, 2021, pp 2015-230. DOI: 10.1093/ oso/ 9780198859307.003.0016. For a recent proposal on how to combine the scenario approach with the Bayesian method in a Bayesian Network, see C. Dahlman and H. Jellema, Scenario-Based Bayesian Networks for Legal Evidence, submitted.

[43] Dahlman 2024.

on this systematization he investigated whether and to what extent training in Bayesian reasoning helps lawyers to recognise these probabilistic thinking errors.[44] Jellema has translated Dahlman's overview of errors in reasoning into Dutch and made it accessible to practising lawyers, among others by providing examples from Dutch case law.[45] We plan to convert this overview into a checklist with 'critical questions' that lawyers can ask when analysing evidence. This checklist should help prevent thinking errors such as the prosecutor's fallacy which was discussed in section 2.

Our research group also investigates whether judges themselves can model court cases as a whole using the Bayesian model. We focus in particular on the phase of the court's deliberation chamber in which, as already suggested in section 4.1, judges arrive at an evidential judgement in a 'traditional' manner and compare this judgement with the outcome of their own Bayesian analysis. In our view, this analysis should consist of setting up a Bayesian Network with the help of a forensic advisor. In terms of the distinctions mentioned above, we focus on the combination of questions 1c, 2b and 3b.

What is a Bayesian Network (BN)? A BN looks like a mind map. However, it is a mind map whose nodes and connections are precisely defined. It is used to visualise and analyse the often complex dependencies between hypotheses and evidence.[46]

When drawing up a BN, the primary focus is not on the quantitative outcome that a BN can deliver, but rather on the fact that the activity of drawing up a BN jointly can support both the individual thought process and the discussion between judges. It should be emphasised that there is no single correct, true or best BN, but rather a BN on which the drafters have agreed or disagreed after consultation and discussion and which they can explain and justify to others.[47] Drawing up a BN can help the drafters to think in a structured way about complex relationships between hypotheses and evidence, making their assumptions transparent and open to discussion and potentially preventing probabilistic errors in reasoning (see Figure 1).

All this might still sound rather abstract, so let me give two examples. Suppose there are two witness statements in a criminal case. Drawing up a BN in combination with the aforementioned checklist of critical questions 'forces' the authors to think about the 'critical question' of whether the two witnesses spoke to each other or made their statements completely independently of each other. The answer to this critical question ('Is the evidence dependent or independent?') is important because

---

[44] C. Dahlman, Training legal fact-finders to recognise probabilistic fallacies, *Law, Probability and Risk* Volume 24, issue 1, 2025, https://doi.org/10.1093/lpr/mgaf010.

[45] Jellema and Dahlman 2025. The use of a checklist is in line with the view of Prakken and Meester 2017, who write: 'The power of Bayesian thinking lies primarily in its therapeutic and pedagogical value: with small examples, it is relatively easy to show that certain seductive patterns of thinking about probabilities are incorrect.' and: 'This rule arms us against errors in reasoning.'

[46] For an accessible introduction to Bayesian Networks, see David A. Lagnado, *Explaining the Evidence,* Cambridge, Cambridge University Press, 2022, chapter 3.

[47] It can be said, however, that some models are logically incorrect.

independent evidence has greater probative force than dependent evidence. This also shows in the different modelling of the witness statements (see Figure 2).

Figure 1. Bayesian Network

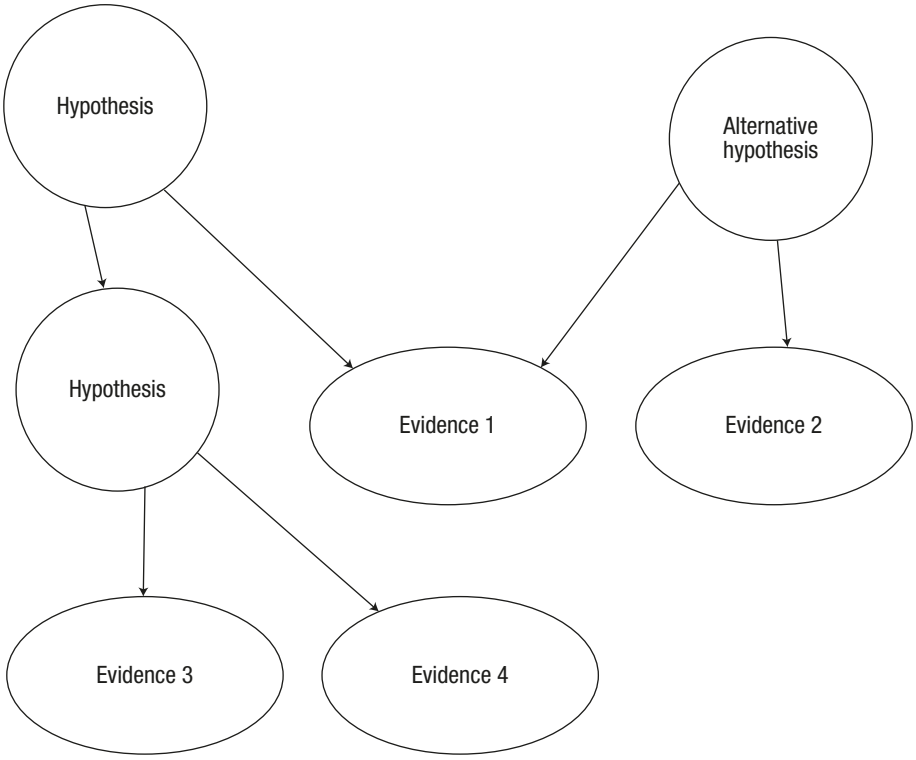
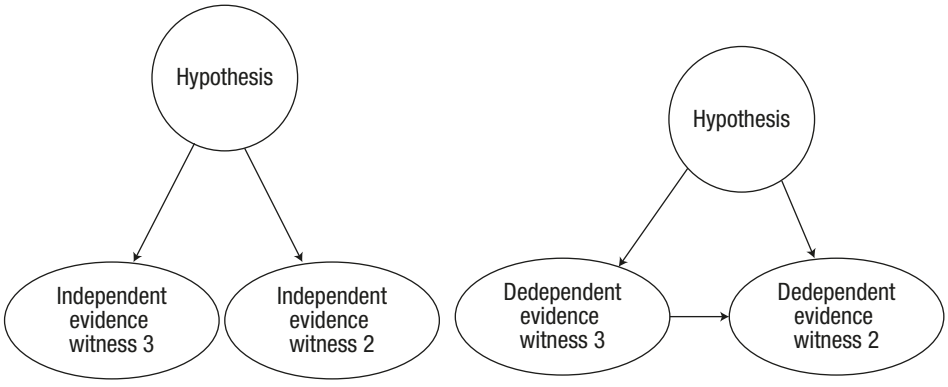
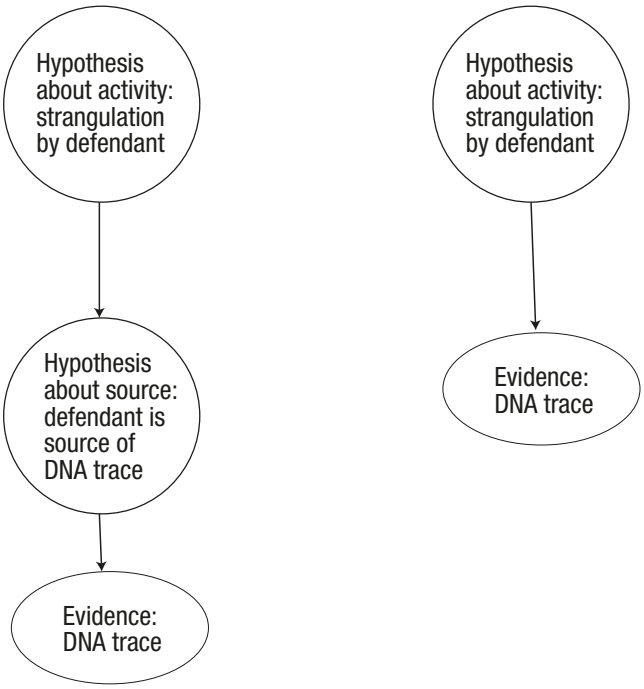


Figure 2. BN's with dependent and independent evidence



Another example: in a court case, the NFI has written a report on DNA evidence. Creating a BN helps the judges to consider the 'critical question' of exactly which question the forensic report answers. Does the report provide a LR about the source of the DNA, i.e. whether the the defendant is the source of the DNA, or does it provide a LR about the activity that caused the DNA to end up at the crime scene, for example through an act related to the crime such as strangulation? This distinction is important because the probative force of DNA at the activity level is often much lower than at the source level. Ignoring this distinction can therefore lead to an overestimation of the probative force. Here too, drawing up a BN in combination with asking the 'critical question' ("To which hypothesis does the LR relate?", "Source or activity?") helps to raise awareness of this crucial distinction which is also visible in the two BNs. [48]

Figure 3. BN's of DNA analysis at source and activity level



[48] Note that a similar error can also occur in the example of shoe prints that was discussed in section 2. The NFI reports on the suspect's shoe, not on the hypothesis that it was the suspect who wore the shoe and made the print. The LR that the suspect made the print is possibly much smaller than the LR that the print was made with his shoe.

Let us return to the use of Bayesian Networks by judges in the court's deliberation chamber. As mentioned above, the primary focus is not on the end product, but on the process of jointly creating a BN. Drawing up a BN makes both choices and errors in reasoning transparent, so that they can be discussed. The creation of a BN ultimately also leads to a product, a network. This network can be qualitative (as in the images above) or quantitative. When creators assign numbers to the prior probabilities and the LRs, the software used to create BNs can perform the often highly complex calculations for the user.[49]

In section 4.2, we saw that the Court of Appeal in 's-Hertogenbosch in its ruling of 5 February 2018 (ECLI:NL:GHSHE:2018:421) argues that quantifying subjective probability estimates gives the calculations an unjustified appearance of objectivity. This has been referred to as 'garbage in, garbage out' (Prakken and Meester, 2017, p. 191). This critique certainly has some merit. It is a legitimate question to ask what value numbers add if they are not scientifically substantiated. The main answer is that judges have to make a probability assessment anyway. Assigning a number to verbal expressions such as "probable" or "very probable" can ensure that judges use these terms in a more uniform manner. Like the NFI, judges too can indicate a range of numbers, for example a lower limit of 10 and an upper limit of 100. Another important function of quantification is that, in practice, it proves difficult to intuitively estimate the effect of combining prior probabilities with several LRs. By quantifying these probabilities and varying the numbers within a range which is deemed reasonable by judges, the creators of the BN can gain insight into the effect of the estimates they make. If the quantitative outcomes of the BN differ significantly from the qualitative outcome of the 'traditional' legal deliberations, this may be a reason to critically compare the 'traditional' evidence reasoning and the BN. But even if there is no reason to change the evidentiary decision, drawing up a BN can contribute to better evidential reasoning because it leads to clarification and careful consideration of choices.[50]

6. CONCLUSION AND RECOMMENDATIONS

The Bayesian model is internationally accepted for estimating the likelihood ratio (LR), i.e. the probative force of forensic evidence. The Court of Appeal in 's-Hertogenbosch adopted this position in its ruling of 12 November 2020 (question 1a). However, it is unclear whether and under what conditions the method is also accepted in Dutch case law for analyses by non-forensic experts such as legal psychologists (question 1b). Bayesian analyses of criminal cases as a whole by an 'omniscient' ex-

[49] Software to create BN's such as Hugin and AgenaRisk is comparable to the software of a calculator that performs calculations for us based on our input. Of course, this quantification step assumes that the judge has quantified the values of priors and evidence.

[50] More on the constitutional importance of sound reasoning based on evidence in Mackor 2024.

pert have not been used been by lower courts (question 2a). However, there is still no clarity about the use of an integral Bayesian analysis drawn up by the court itself (question 2b). It is to be hoped that the Dutch Supreme Court will provide clarity on these topics.

I have argued that the lower courts have rightly rejected the modelling of criminal cases as a whole by experts. The main argument is not that 'the method is still too uncertain', but that this would allow experts to sit on the seat of the judge, which is unacceptable from a rule of law perspective. However, we must avoid throwing the Bayesian baby out with the bathwater. Further research is needed to determine whether and how the Bayesian model can help judges. In particular we want to investigate whether the use of a checklist of 'critical questions' can help prevent probabilistic errors in reasoning and thus improve the quality of judicial evidence evaluation. In addition, it must be investigated whether the creation of Bayesian Networks by judges in the court's deliberation chamber can improve the quality of judicial evidence evaluation. First, it is important to investigate whether it is feasible for judges to create a BN themselves. Feasibility in this context concerns both the time it takes to create a BN and whether judges are capable of creating a sound BN on their own or in collaboration with a forensic advisor. It is also, from a rule of law perspective, necessary to investigate the influence of the forensic advisor on the modelling. Theoretical and empirical research is needed to answer these questions. Over the next three years, we hope to be able to provide answers to these and other questions as part of our NWO research project Preventing Miscarriages of Justice.[51] [52]

[51] See note *. For recent work and work in progress of our research group, see Thomas Boersma, Bayesiaanse netwerken in bewijsminimumzaken: toepassing op de zaak Ali B [Bayesian Networks in minimal evidence cases: an application to the case of Ali B], in progress and Thomas Boersma, Logische en correcte bewijswaardering: Bayesiaans redeneren in zaken met seksueel grensoverschrijdend gedrag, [Logical and correct evidence evaluation: Bayesian reasoning in cases of sexual transgressive behaviour], Tijdschrift Recht en Arbeid, forthcoming (the first Dutch Bayesian article on labour law). C. Dahlman and H. Jellema, submitted, on combining Bayesian and scenario approaches. Leya Hampson and Ludi van Leeuwen, Investigating the value of qualitative Bayesian networks of complete cases as "double-check" tools on traditional judicial reasoning: An exploratory study, forthcoming in the *Proceedings of the AI4EVIR Workshop on AI for evidential reasoning,* December 9, 2025, Turin, Italy. A.R. Mackor and H. Prakken On Reporting Likelihood Ratio's of Exhaustive and Non-Exhaustive Hypotheses about Rare Events in Criminal Case, forthcoming in the *Proceedings of the AI4EVIR Workshop on AI for evidential reasoning, December 9, 2025, Turin, Italy*.

[52] For other relevant recent research on BN's see Nikolay Babakov, Ehud Reiter and Alberto Bugarín, Explaining Bayesian Networks Reasoning to the General Public: Insights from the User Study, *Multimodal, Affective and Interactive eXplainable AI*, 2025 for a large user study (124 participants) evaluating the interpretability of BN. In our research group, however, we do not only investigate the passive understanding of BN's, but focus on active co-creation of BN's.

REFERENCES

Alkemade, F. (2018). Makkers, staakt uw wild geraas. Een oproep tot de Nederlandse deskundigen in forensic reasoning om een consensus te formuleren. *Expertise en Recht*, *2018*(1), 26-35.

Babakov, N., Reiter, E., & Bugarín, A. (2025). Explaining Bayesian Networks Reasoning to the General Public: Insights from the User Study. *Multimodal, Affective and Interactive eXplainable AI*.

Boersma, T. (forthcoming). Bayesiaanse netwerken in bewijsminimumzaken: toepassing op de zaak Ali B.

Boersma, T. (forthcoming). Logische en correcte bewijswaardering: Bayesiaans redeneren in zaken met seksueel grensoverschrijdend gedrag. *Tijdschrift Recht en Arbeid*.

Dahlman, C. (2024). A systematic account of probabilistic fallacies in legal fact-finding. *The International Journal of Evidence & Proof*, *2024*(1), 45-64.

Dahlman, C. (2025). Training legal fact-finders to recognise probabilistic fallacies. *Law, Probability and Risk*, *24*(1). *https://doi.org/10.1093/lpr/mgaf010*

Dahlman, C., & Jellema, H. (forthcoming). Scenario-Based Bayesian Networks for Legal Evidence.

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.

De Roo, R., Remijn, W., Meeuwissen, J., Kruithof-van Esch, J., Van der Heijden, S., Claushuis, M., & Van Blijswijk-Kieftenbeld, L. (2021). Forensisch adviseurs actief bij alle gerechten in Nederland: tijd voor een terugblik. *Expertise en Recht*, *2021*(6), 234-239.

Gerechtshof Amsterdam [Tribunal de Apelación de Ámsterdam]. (2021, 12 de abril). ECLI:NL:GHAMS:2021:986.

Gerechtshof Den Bosch [Tribunal de Apelación de 's-Hertogenbosch]. (2008, 24 de septiembre). ECLI:NL:GHSHE:2008:BF2188.

Gerechtshof Den Bosch [Tribunal de Apelación de 's-Hertogenbosch]. (2016, 22 de noviembre). ECLI:NL:GHSHE:2016:5165.

Gerechtshof Den Bosch [Tribunal de Apelación de 's-Hertogenbosch]. (2017, 30 de mayo). ECLI:NL:GHSHE:2017:2300.

Gerechtshof Den Bosch [Tribunal de Apelación de 's-Hertogenbosch]. (2018, 5 de febrero). ECLI:NL:GHSHE:2018:421.

Gerechtshof Den Bosch [Tribunal de Apelación de 's-Hertogenbosch]. (2018, 9 de octubre). ECLI:NL:GHSHE:2018:4181.

Gerechtshof Den Bosch [Tribunal de Apelación de 's-Hertogenbosch]. (2020, 12 de noviembre). ECLI:NL:GHSHE:2020:3483.

Gerechtshof Den Haag [Tribunal de Apelación de La Haya]. (2015, 14 de octubre). ECLI:NL:GHDHA:2015:2860.

Hampson, L., & Van Leeuwen, L. (en prensa). 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.

Hoge Raad [Tribunal Supremo de los Países Bajos]. (2020, 25 de agosto). ECLI:NL:HR:2020:1317.

Hoge Raad [Tribunal Supremo de los Países Bajos]. (2021, 15 de junio). ECLI:NL:HR:2021:788.

Jellema, H., & Dahlman, C. (2025). Probabilistische drogredenen in de rechtszaal: een systematisch overzicht. *Expertise en Recht*, *2025*(3), 64-74.

Koninklijk Nederlands Meteorologisch Instituut [KNMI]. (s.f.). *Het regent niet vaker, wel harder*. *https://www.knmi.nl/over-het-knmi/nieuws/het-regent-niet-vaker-wel-harder*

Lagnado, D. A. (2022). *Explaining the evidence*. Cambridge University Press.

Mackor, A. R. (2024). Risks of incorrect use of probabilities in court and what to do about them. En A. Placani & S. Broadhead (Eds.), *Risk and responsibility in context* (pp. 94-108). Routledge.

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.

Mackor, A. R., & Van Koppen, P. J. (2021). The scenario theory about evidence in criminal law. En C. Dahlman, A. Stein, & G. Tuzet (Eds.), *Philosophical foundations of evidence law* (pp. 215-230). Oxford University Press. *https://doi.org/10.1093/oso/9780198859307.003.0016*

Meester, R., & Stevens, L. (2021). Correct redeneren: wat een Bayesiaanse analyse zegt over vier recente strafrechtelijke uitspraken. *Expertise en Recht*, *2021*(3), 112-121.

Nederlands Forensisch Instituut [NFI]. (2017). *Vakbijlage Waarschijnlijkheidstermen*. *https://www.forensischinstituut.nl/publicaties/publicaties/2017/10/18/vakbijlage-waarschijnlijkheidstermen*

Niessen, R. E. C. M. (Advocaat-Generaal). (2008, 14 de marzo). Conclusión ECLI:NL:PHR:2008:AU0837.

Parket bij de Hoge Raad [PHR]. (2020, 16 de junio). Conclusión ECLI:NL:PHR:2020:560.

Parket bij de Hoge Raad [PHR]. (2021, 20 de abril). Conclusión ECLI:NL:PHR:2021:391.

Parket bij de Hoge Raad [PHR]. (2025, 25 de mayo). Conclusión ECLI:NL:PHR:2025:378.

Prakken, H. (2018). Kansoordelen door deskundigen: over "logisch" rapporteren en wat daarbij mis kan gaan. *Ars Aequi*, *67*, 740-747.

Prakken, H. (2020). Rechters en Raadsheren over Bayes. En R. Horselenberg, V. van Koppen, & J. de Keijser (Eds.), *Bakens in de Rechtspsychologie: Liber Amicorum voor Peter van Koppen* (pp. 287-300). Boom Uitgevers.

Prakken, H., Bex, F. J., & Mackor, A. R. (2020). Editors' Review and Introduction: Models of Rational Proof in Criminal Law. *Topics in Cognitive Science*, *12*, 1053-1067.

Prakken, H., & Mackor, A. R. (en prensa). On reporting likelihood ratios of exhaustive and non-exhaustive hypotheses about rare events in criminal case. *Proceedings of the AI4EVIR Workshop on AI for Evidential Reasoning*, Turín, Italia.

Prakken, H., & Meester, R. (2017). Bayesiaanse analyses van complexe strafzaken door deskundigen. Betrouwbaar en zo ja: nuttig? *Expertise en Recht*, *2017*(5), 185-197.

Rassin, E., Arbiyah, N., Boskovic, I., Otgaar, H., & Merckelbach, H. (2022). Likelihood ratios in psychological expert opinion, and their reception by professional judges. *The International Journal of Evidence & Proof*, *26*(4), 325-341. *https://doi.org/10.1177/13657127221119545*

Rechtbank Gelderland [Tribunal de Distrito de Güeldres]. (2016, 16 de febrero). ECLI:NL:RBGEL:2016:825.

Rechtbank Limburg [Tribunal de Distrito de Limburgo]. (2015, 25 de febrero). ECLI:NL:RBLIM:2015:1588.

Rechtbank Oost-Brabant [Tribunal de Distrito de Brabante Oriental]. (2017, 19 de diciembre). ECLI:NL:RBOBR:2017:6556.

Rechtbank Zeeland-West-Brabant [Tribunal de Distrito de Zelanda-Brabante Occidental]. (2016, 23 de mayo). ECLI:NL:RBZWB:2016:3060.

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.

www.quaestiofacti.com