

POR QUÉ UN ESTÁNDAR DE PRUEBA SUBJETIVO Y AMBIGUO NO ES UN ESTÁNDAR

Larry Laudan

Unam, México

RESUMEN. Un estándar intrínsecamente subjetivo no nos ofrece protección alguna contra los riesgos de las decisiones arbitrarias. En un sistema con tales estándares, pruebas que pueden convencerle a A de la culpabilidad de X no van a persuadirle a B de que X es culpable. Obviamente, no podemos esperar ni uniformidad ni justicia si cada juez utiliza sus propios criterios. En vez de utilizar estándares de prueba subjetivos, como la convicción íntima o la ausencia de una duda razonable, propongo aquí varios estándares objetivos y exigentes. Ellos son objetivos porque no dependen de los caprichos de si o no cierto juez tenga una convicción firme sobre la culpabilidad del acusado. Son exigentes porque no queremos condenar a alguien excepto que tengamos un prueba muy sólida y robusta.

ABSTRACT. An intrinsically subjective standard offers us no protection against the risks of arbitrary decisions. In a system with standards such as these, evidence which can convince A of X's guilt will not convince B that X is guilty. Obviously, we can expect neither uniformity nor justice if each judge uses his own criteria. Instead of using subjective standards of proof, such as intimate conviction or the absence of reasonable doubt, I thus put forward a number of objective and demanding standards. They are objective because they do not depend on the whims of a judge, on whether or not a certain judge is strongly convinced of the accused's guilt. They are demanding because we do not wish to find anybody guilty unless we have strong and solid evidence.

Una introducción meta-filosófica

No soy ni abogado ni filósofo del derecho. Lo que hago es epistemología. Inevitablemente, esta confesión requiere una *defensa* (o por lo menos una etiqueta de advertencia) porque, con excesiva frecuencia, ser epistemólogo significa estar preocupado por derrotar el escepticismo filosófico o por establecer si es el caso que ‘John sabe que X’ supone que ‘John sabe que John sabe que X’. Seguramente quedarán aliviados al saber que estos temas *no son* los míos. Aunque algo de ese alivio puede desaparecer al descubrir que soy igualmente indiferente a preocupaciones del estilo de las de Kelsen y Hart respecto del concepto del derecho o en referencia al tema de las fuentes legítimas de la autoridad jurídica.

Mis intereses son más modestos y menos ‘*a priori*’ que aquellos más familiares para la epistemología analítica o para la filosofía del derecho. Estoy interesado, más bien, en lo que es, en última instancia, una pregunta empírica, a saber: ¿cuán bien (o cuán mal) hacen su función los sistemas de investigación existentes?—entendidos en un sentido amplio como motores para descubrir la verdad sobre temas significativos relacionados con los seres humanos. Y, deslizándome rápidamente a lo normativo, ¿cómo podrían ser mejoradas estas prácticas para reducir el error? Durante mucho tiempo mi principal interés fue averiguar cómo funcionan las ciencias naturales en tanto que un sistema para generar el conocimiento del mundo natural. Posteriormente, mis intereses giraron hacia el derecho penal.

Comienzo con la premisa, quizás de carácter contrafáctico, de que *el objetivo principal de un juicio penal es averiguar la verdad acerca de la comisión de un supuesto delito*. Mi programa de investigación —aunque este término pueda resultar demasiado ambicioso para mis esfuerzos preliminares en este área— es explicar cómo podría estructurarse un juicio —en términos de reglas de prueba y de procedimiento— para elevar al máximo la probabilidad de que el resultado sea un fallo verdadero. Por supuesto, deseáramos también que el resultado de un juicio sea justo. No diré aquí nada acerca de esta segunda meta, salvo observar que la veracidad de los resultados es condición necesaria (aunque no siempre suficiente) de su justicia. En resumen, mis intereses se centran en lo que llamo *epistemología jurídica*.

Dado que este término no coincide exactamente con el del uso generalizado —en realidad es apenas el mismo— quizás sea útil empezar diciendo brevemente algo acerca de cómo concibo esta área de investigación que está surgiendo. Cada juicio y, más generalmente, cada sistema de justicia penal, está gobernado por tres tipos de valores: 1) Muchos de esos valores son extra-epistémicos. Entre estos podemos destacar las consideraciones de oportunidad, las relacionadas con los derechos de los acusados, los vinculados

con la transparencia y el debido proceso, y cosas por el estilo. No diré nada aquí en referencia a estos valores en tanto que ellos quedan fuera del ámbito de la epistemología jurídica. 2) Un segundo conjunto de valores que dirigen los juicios penales y las investigaciones constituyen lo que llamo el *núcleo duro de la epistemología jurídica*. El interés en éste ámbito está en reducir o aminorar la probabilidad de un juicio erróneo (donde por ‘erróneo’ se entiende específicamente ‘falso’). Los errores principales son la condena de alguien que no cometió el delito o la absolución de aquél que sí lo hizo. El núcleo duro de la epistemología jurídica se interesa precisamente en cómo hacer para que éstos errores sean tan improbables como permita la evidencia. 3). Hay un tercer conjunto de valores, por lo menos tan importante como los otros dos, que son *quasi-epistémicos*. A falta de un mejor término, los llamo el *núcleo débil de la epistemología jurídica*. Aquí, el interés no está en la reducción de los errores sino en su *distribución* de una manera particular. Llamo a estos valores quasi-epistémicos porque, aunque ellos se centran (al igual que los valores del núcleo duro) en el control del error, sus motivos no derivan de su relación con la verdad ni con la reducción del error sino de una decisión política según la cual cierto tipo de errores es peor, menos aceptable, que otros. Aunque el epistemólogo no puede decirle al sistema jurídico cómo deben ser distribuidos estos errores, lo que sí puede hacer es ayudarlo –una vez se ha decidido el peso relativo socialmente aceptable de los errores– preocupándose por cómo puede ser satisfecho de una mejor manera este peso relativo (por ejemplo, a través de la determinación de la altura de un estándar de prueba). Muchas de las doctrinas más tradicionales o de los preceptos del derecho penal caen en este ámbito: la presunción de inocencia, la doctrina del *in dubio pro reo*, la idea de que es la acusación quién soporta totalmente la carga de la prueba, y –la más importante– el estándar de prueba.

Este artículo se sitúa sin ambigüedades en algunos temas de esta tercera categoría. Presentaré una pregunta –o mejor dicho, una serie de preguntas– acerca del núcleo débil de la epistemología jurídica. Y, aunque mi interés no estará vinculado con la reducción de los errores, al considerar los mecanismos que quizás tendríamos que utilizar para distribuirlos en la forma en que deseamos, inevitablemente, el tema de la verdad acerca de los delitos surgirá durante el análisis. Sin embargo, cualquier progreso que sea capaz de hacer en este punto no afectará al núcleo duro, esto es, a las preguntas acerca de la reducción del error y de su minimización. Éstas, las dejaré para otra oportunidad.

En concreto, quiero tratar directamente en este ensayo las siguientes cuestiones: qué es un estándar de prueba (en adelante: EdP), cómo funciona, y a qué podría parecerse un EdP conceptualmente sólido. Hacia el

final de mi trabajo, también me ocuparé de decir algo acerca de cómo este particular mecanismo para la distribución de los errores –EdP– se relaciona con un conjunto de conceptos familiares (el beneficio de la duda, la presunción de inocencia, y la carga de la prueba) los cuales, al igual que el EdP, son también instrumentos diseñados para distribuir el error antes que para reducirlo.

Las dos rutas convencionales para un EdP: la duda razonable (alias: la convicción íntima), y las probabilidades bayesianas

En este artículo, quiero recoger algunos aspectos muy generales acerca del EdP en el ámbito específico de los juicios penales. El análisis que aquí presento surge de mis esfuerzos por entender cómo funcionan estos estándares (esto es, lo que presupone, lo que traen consigo, y cómo se aplican) en la ley Anglosajona, en primer lugar, y, por razones biográficas obvias, en la ley mexicana en segundo lugar. Si la crítica que ofreceré puede ser aplicada en otras partes es algo que no estoy, por ahora, dispuesto a afirmar. Supondré, para avanzar en el tema, que el asunto que pretendo refutar se aplica de forma similar en la mayoría de los sistemas europeos de justicia penal, aunque estoy abierto a la corrección si es que he lanzado mi red demasiado lejos.

Si observamos el sistema anglosajón o el sistema continental hay, esencialmente, dos clases de discursos bastante bien definidas acerca del EdP. La primera se desarrolla principalmente en la sala, entre jueces y jurados, y en las decisiones de los tribunales de apelación. Aquí, la pauta general es configurar el EdP en términos cualitativos: desde este punto de vista, aquél a quien corresponda juzgar los hechos deberá condenar al acusado sólo en el caso en que concurra una “certeza moral” o una “certeza que va más allá de la duda razonable” de que el acusado cometió el delito. El juez espera tener una “convicción íntima” acerca de esa culpabilidad, o (como en México) tener lo que él considera una “prueba plena” de la culpa de acusado.

El otro discurso acerca del EdP tiende a ocurrir entre juristas académicos –anglosajones y continentales– cuya preferencia decidida es hablar de los EdP como si fueran (bayesianas) probabilidades posteriores. En este segundo caso, la idea es que el EdP establece un umbral de probabilidad (generalmente entre el 90% y el 95%) que debe lograrse antes de que el juzgador de los hechos esté justificado en condenar al acusado.

Estos dos discursos, el práctico y el académico, no están enteramente separados. El que llamaré “estándar de la prueba académico” ha delineado típicamente el trabajo estándar del juez sobre la base de alguna afirmación según la cual la prueba más allá de una duda razonable (en adelante: BARD)

debería ser entendida como el equivalente de un EdP del 90%¹. Pero mientras que los académicos están felices con intentar salvar la distancia entre los dos discursos, los jueces tienden a ser reacios prefiriendo formular el EdP exclusivamente en términos no-cuantitativos².

No es mi objetivo procurar reconciliar estas tensiones entre un estándar cualitativo y uno cuantitativo (aunque mis propias preferencias en este asunto quedarán aclaradas en las observaciones que seguirán). Esto sería un objetivo digno de conseguir si uno o ambos de estos enfoques fueran siquiera medianamente prometedores. Pero como no es así, no lo haré. En vez de eso, sugeriré que ambos discursos se confunden sistemáticamente acerca de lo que es un EdP y acerca del trabajo que podemos esperar que un EdP haga por nosotros. Para provocar la discusión, iré tan lejos como para sugerir que los jueces y los juristas académicos no tienen más que una comprensión meramente superficial de lo que es un EdP y de por qué lo necesitamos.

La subjetividad y la ambigüedad de los estándares actuales

Los jueces americanos e inglés son fetichistas acerca de lo que les gusta llamar “confianza subjetiva”. Por supuesto, oficialmente, el estándar anglosajón es el de la prueba más allá de una duda razonable antes que el de sólo una alta confianza. Pero, como he mostrado en otra parte³, el estándar de la duda razonable ha sido cambiado más allá de lo reconocible desde su creación hace ya dos siglos. Aquello que una vez significó certeza moral y fue relacionado con la epistemología de Locke, ha sido deliberada y lentamente desplumado por la jurisprudencia americana durante el siglo XX y ahora ‘más allá de una duda razonable’ ha quedado reducido simplemente a un ‘carecer de toda duda’, esto es, a tener una alta confianza subjetiva. Hoy en día, el punto de vista oficial de la Corte Suprema es que una condición *necesaria y suficiente* para condenar al acusado en un juicio penal es una creencia firme de la culpabilidad por parte del jurado. Al creer (como los juristas del sistema continental) en una teoría de la “libre prueba”, los jueces no hacen nada para indicar al jurado qué tipos de inferencias, o qué clases de evidencias, justificarían la formación de una creencia firme en la cul-

¹ Lawrence Tribe es una notable excepción entre los juristas académicos por su ardua resistencia a cualquier sugerencia de que el estándar de prueba puede ser intercambiado por una probabilidad.

² Hay importantes razones políticas, que yo no seguiré aquí, por las que los jueces tienden a ser especialmente resistente a la sugerencia que el EdP es una probabilidad: la más importante de ellas es que tal declaración sería una admisión explícita de que el sistema está diseñado para cometer errores.

³ Véase LAUDAN, Larry; *Is Reasonable Doubt Reasonable?* en 9 LEGAL THEORY 295 (2003).

pabilidad del acusado. Aparentemente, todo lo que importa es si, habiendo oído toda la evidencia, la creencia de los jurados es firme y sólida. Y si no es así, entonces debe votar la absolución. Los jueces ni tan siquiera explican a los jurados lo que significa tener una duda razonable, manteniendo de manera vergonzosa que esta noción es auto-evidente. Esta tesis ha sido constantemente refutada por los numerosos pedidos de los jurados a los jueces para que expliquen lo que significa esta doctrina. Básicamente, en estas circunstancias, la duda razonable queda reducida a cualquier duda que un jurado quiera utilizar para absolver a alguien. Bajo las circunstancias es difícil decidir cuál es lo peor: la subjetividad del estándar actual o su falta de una definición clara.

La prueba de esta tesis requiere mucho más espacio del que dispongo en este trabajo. Pero, dado que ya he desarrollado este tema en otra parte, remitiré a los lectores curiosos o escépticos a mi propuesta. Simplemente la presupondré en las consideraciones que aquí realizaré sobre el lamentable estado de la discusión contemporánea de la prueba más allá de una duda razonable entre nosotros los anglosajones. (Es sintomático de la desesperación que rige la situación actual del sistema del *common law* en Inglaterra el reciente abandono del reclamo para que los jurados utilicen el estándar razonable de la duda y su reemplazo de manera oficial por el estándar de la firme convicción. Esto no es tanto un cambio en la práctica como un reconocimiento explícito de lo que la práctica existente ha llegado a ser.)

Por su parte, la versión académica del EdP anglosajón es igual de profundamente subjetiva. Supone que el jurado empieza un juicio asignando una cierta (generalmente inespecífica pero tendencialmente baja) probabilidad previa respecto de la culpabilidad del acusado. Cuando oye la evidencia y el testimonio, el jurado revisa constantemente su estimación de la probabilidad (probabilidad posterior) de culpabilidad del acusado. Al final del juicio determina si su, hasta ahora múltiples veces revisada, estimación alcanza el número especificado por el EdP (típicamente, alrededor de 0,9 o 0,95). Si este fuese el caso, entonces deberá condenar; si no, deberá absolver. Sostengo que este estándar es profundamente subjetivo porque la estimación final de la presunta culpabilidad dependerá inevitablemente, entre otras cosas, de las corazonadas subjetivas *iniciales* del jurado acerca de la culpabilidad o inocencia⁴. Los jurados que comiencen el juicio con prioridades diferentes (como son llamadas), aunque sean confrontados exactamente con la misma

⁴ Para un excelente análisis de muchos de los problemas asociados con la aplicación de las probabilidades clásicas a los tipos de inferencias apropiadas a un acto jurídico véase TARUFFO, Michele, *La Prueba de los hechos*, Ed. Trota, Madrid, 2002 y COHEN, Jonathan, *Lo probable y lo demostrable* (1991).

evidencia, alcanzarán estimaciones diferentes sobre la probabilidad final de la culpa. (La modelización computacional de ciertos juicios típicos hecha, entre otros, por Robert Causey, sugiere (a) que las estimaciones hechas por jurados generalmente no convergen al final y (b) que prácticamente ninguna evidencia concebible puede desplazar una evaluación previa muy baja de la culpabilidad hacia un nivel próximo al 0,9. Así, si un jurado toma la presunción de inocencia seriamente y empieza el juicio asumiendo que la probabilidad inicial de inocencia es (digamos) 0,95, no hay virtualmente evidencia alguna que pueda cambiar plausiblemente esa estimación a 0,9 o 0,95 a favor de la culpabilidad.)

Pero, incluso si ignorásemos la cuestión de la subjetividad de las prioridades, aún hay problemas más graves frente al punto de vista de que un EdP debe entenderse en términos probabilísticas. Esto sugiere que el discurso técnico de las probabilidades no es seguramente el mejor medio para formular un EdP. Hay varias razones para pensar esto. Me centraré en dos de ellas en particular, aunque hay muchas otras (por ejemplo la demostración de Cohen de que el teorema de multiplicación del cálculo de la probabilidad no tiene sentido de la manera en que se utiliza en los juicios penales). Cualquiera de estos dos problemas, tomado aisladamente, es suficiente como para generar importantes dudas acerca de la conversión del EdP en probabilidades. Tomados conjuntamente, pienso que socavan completamente este proyecto.

Estimando las probabilidades de una creencia

El supuesto más apropiado, incluso paradigmático, de asignación de probabilidades a unas creencias surge en contextos de acontecimientos de carácter específicamente estadístico. Si sé que el catálogo de la *American Legion* organiza una rifa de un tanque excedente de la Segunda Guerra Mundial, y sé también que la rifa consistirá en 1.000 boletos, puedo decidir comprar 900. Suponiendo que la rifa sea justa, creeré que soy el probable ganador del tanque. Incluso más, seré capaz de informar de que estoy un 90% seguro de la verdad de la hipótesis según la cual pronto seré el dueño de un tanque con rancio abolengo. Aquí, la determinación cuantitativa de mi grado de confianza es una cuestión rápida dentro-de-la-envoltura del cálculo. Ahora bien, cuando nos desplazamos desde este específico tipo de hechos hacia consideraciones más amplias resulta más difícil asignar las probabilidades concretas de una creencia.

Suponga que soy jurado en un juicio por robo con arma. He escuchado muchas pruebas inculpatorias y sólo alguna exculpatoria. Habiendo oído con atención las pruebas, y habiendo pensado seriamente sobre ellas, he llegado a la conclusión de que el acusado cometió el delito. Ahora tengo que decidir si voto condenándolo o absolviéndolo. Sé que mi creencia de que el

acusado lo hizo no es, por sí misma, suficiente para decidir sobre este asunto ya que el juez me ha informado de que la mera preponderancia de pruebas no es suficiente. Podría ocurrírseme intentar determinar cuál es la probabilidad que tiene realmente la hipótesis de culpabilidad del acusado. De hecho, si el EdP del delito fue definido explícitamente en términos de probabilidades, estaré obligado a hacer tal determinación. Esto, en general, no es algo que ni yo ni la mayoría de los jurados podamos hacer con algún grado de certeza. Aunque quizás pueda decir que creo firmemente en la culpabilidad del acusado, que encontré la presentación del caso por parte de la acusación bien construida mientras que el de la defensa me pareció débil, estaría en apuros si tuviese que asignar una probabilidad de 95% o 85% o 75% a la hipótesis de la acusación conforme a la cual el acusado cometió el delito. E incluso si me sacara del sombrero algún número en concreto, se requeriría algún elemento más para poder imaginar que éste representa exactamente algo así como mi grado de creencia en la culpabilidad del acusado.

No estoy solo en esto. Si analizamos fuera del ámbito del derecho penal otras formas de investigación empírica, encontramos reticencias similares de los investigadores para asignar probabilidades específicas a sus creencias. Prácticamente todos los físicos en el mundo moderno creen que la teoría general de la relatividad de Einstein es probablemente verdadera. Sin embargo, me aventuro a pensar que *no* hay físico que pueda decirle qué probabilidad específica le asigna a la hipótesis de Einstein. Más aún, probablemente un físico estaría en apuros si tuviera que indicarle aunque sólo sea el pequeño intervalo en el que se encuentra dicha probabilidad. Por ejemplo, si usted le preguntara: “¿es la probabilidad de la teoría general mayor al 95%?” la inmensa mayoría de físicos se encogerían de hombros y le dirían que no están seguros. Pues bien, si una teoría tan bien fundada como la de Einstein, una teoría respecto de la cual los científicos han estado pensando y probando arduamente durante un siglo, goza de un amplio apoyo que no es cuantificable, ¿qué razón podemos tener para pensar que un jurado, que investigó una hipótesis en unas pocas horas (en un juicio típico) o en una pocas semanas (en uno atípico) está situado en una mejor posición que la comunidad física para decidir qué probabilidad específica le asignarían a la culpabilidad del acusado en un juicio? El eminente jurista americano, Learned Hand, afirmó que no podría decir cuál es la diferencia entre BARD (más allá de la duda razonable) y la preponderancia de la evidencia⁵. No es difícil de imaginar *su* reacción si hubiese sido conminado a asignar la probabilidad específica en el veredicto de un juicio.

⁵ Véase UNDERWOOD, Barbara D., *The Thumb on the Scales of Justice: Burdens of Persuasion in Criminal Cases*, 86 YALE L.J. 1299, at 1309 (1977).

Los aficionados a la teoría de la probabilidad nos dirán que tales dificultades pueden resolverse con bastante facilidad. Pregunte simplemente a un jurado cuánto estaría dispuesto a apostar y con qué margen a que la hipótesis de la culpabilidad es verdadera. Si, según nos dicen los expertos, él apostaría \$20 o más con la perspectiva de ganancia de \$21 si tiene razón (perdiendo \$20 si está equivocado), entonces está asignando una probabilidad cercana al 95% a la hipótesis de la culpabilidad. Sin embargo, aunque esta propuesta pueda parecer suficientemente plausible, en realidad lo único que hace es desplazar el problema de nivel. Si soy genuinamente incapaz de asignar una probabilidad a mi creencia de que el acusado es culpable, entonces nada podré decir cuando me pidan que haga una apuesta acerca de su culpabilidad.

En resumen, si el estándar de prueba fue formulado en términos de una probabilidad específica, digamos el 90% o el 95%, será casi seguro que muchos de los jurados que han decidido que el acusado cometió el delito estarán en apuros para determinar, de una manera que no sea meramente arbitraria, si su confianza en la culpabilidad satisfizo o no las demandas impuestas por ese EdP.

Pero incluso si este difícil problema pudiera resolverse, aún quedaría uno más grave acechando bajo la superficie. Este segundo problema está dirigido al núcleo de la pregunta de lo que es un EdP y de aquello que se pretende conseguir. Además, este segundo problema contagia no sólo a los estándares probabilísticos de la prueba sino también a muchos estándares no probabilísticos tales como el BARD o el de la convicción íntima. El problema en cuestión es que todos estos estándares de prueba no nos dicen nada acerca de lo que debe contar como una prueba de la culpabilidad. Sea que se especifique el estándar como BARD, como íntima convicción o como algún nivel particular de probabilidad, todos descansan en la misma equivocación. Lo que éstos les dicen al jurado o a otros juzgadores de los hechos es lo siguiente: escucha con atención el testimonio y otras evidencias en el juicio. Cuando el juicio avance tus creencias acerca de los hechos de este caso cambiarán indudablemente de varias maneras, algunas veces de forma sutil y otras de forma drástica. Al final del juicio, tendrás probablemente alguna creencia acerca de la comisión o no del delito por parte del acusado. Si crees que él no cometió el crimen, entonces por supuesto debes absolverlo. Pero, si crees que él lo cometió, deberás hacerte una pregunta adicional. Si el estándar en ese lugar es el BARD, la pregunta se reduce a esto: ¿estás completamente persuadido de que él lo cometió? Si, por el contrario, han sido los abogados académicos los que han impuesto su forma de entender el estándar y han establecido que éste es del 90%, la pregunta que debes hacerte será: ¿tu evaluación de la probabilidad de culpabilidad es mayor del 90%? Si el jurado puede contestar afirmativamente cualquier de estas preguntas,

debe condenar al acusado. En otro caso, deberá absolverlo. Para avanzar con el argumento, concederé al subjetivista que los jurados pueden decir si están firmemente persuadidos o si han asignado una probabilidad mayor del 90% a la hipótesis de la culpabilidad. Aún así, esto sería en vano.

La sencilla familiaridad de esta imagen hace difícil dar un par de pasos hacia atrás y reconocer lo extraña que es esta historia. Y nada de esto resulta más peculiar que el hecho de ofrecer un EdP que no está *fundamentado-libremente*. En vez de eso, el sistema lo único que les dice a los juzgadores de los hechos es lo siguiente: al final del juicio, piense detenidamente acerca de la evidencia; observe cuál es el nivel de confianza que ha alcanzado acerca de la culpabilidad del acusado; y, si usted está realmente seguro de que él cometió el crimen, entonces debe condenarlo; si, por el contrario, usted está algo menos que profundamente persuadido, absuélvalo.

No me andaré con rodeos, esto es una parodia de un 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)– tiene la intención de indicar al investigador o aquel que se está cuestionando cuándo está autorizado a considerar algo 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. En el derecho penal, por el contrario, esta cuestión es ignorada enteramente o sorteada descaradamente mediante algunos trucos. En vez de especificar que el nivel de confianza del jurado respecto de la culpabilidad debe depender del ofrecimiento de una prueba firme, el derecho penal hace que el EdP sea parasitario del nivel de confianza que el investigador (esto es el jurado) tiene respecto de la culpabilidad del acusado. Tenemos una prueba, dice la ley, con tal que los jurados estén totalmente persuadidos de la culpabilidad del acusado o con tal que ellos le asignen una probabilidad mayor que X , en donde X es el EdP probabilístico. En cualquier caso, cuando el jurado haya llegado a un algo grado de confianza, tendremos una prueba. Esto es poner las cosas precisamente al revés.

En caso de que no resulte claro por qué esto pervierte el orden lógico de las cosas, imagínese que les dijésemos a los matemáticos que, de ahora en adelante, contarán con una prueba legítima de un teorema tan pronto como estén convencidos de la verdad de dicho teorema. O, supóngase que les decimos a los epistemólogos que si están sumamente seguros de la existencia de una conexión causal entre A y B , entonces tienen una prueba de ello. Propuestas como estas respecto de la prueba serían objeto de risa. Uno no debería decirle a cualquier juzgador de los hechos: Usted ha proporcionado una prueba de A cuando está firmemente convencido de A . (Ni, tampoco: “Usted tiene una prueba de A con tal que su confianza subjetiva sea más

alta que X.”) Al contrario, lo que le decimos es: Usted no está autorizado para estar totalmente convencido de A a menos que y hasta que usted tenga una prueba de A, añadiendo para completar la propuesta que sus firmes convicciones acerca de A no cuentan en absoluto como si no tuviese una prueba de A. Y entonces procedemos a decirle como sería una prueba de A. Esto es lo que supone tener un EdP. Un EdP apropiado no depende de una confianza subjetiva en una hipótesis; al contrario, el EdP nos indica cuándo la confianza subjetiva está justificada.

En el contraste hasta aquí resaltado, las prácticas probatorias en el derecho penal hacen que la existencia de una prueba de culpabilidad sea parasitaria de la existencia previa de una convicción firme acerca de dicha culpabilidad. El lenguaje usado para dar cuenta de esta convicción subjetiva varía a través de las épocas y de las jurisdicciones. En algunas épocas, ha sido llamada una creencia duradera en la culpabilidad. En otras, se ha considerado una certeza moral, una creencia consolidada, una conciencia satisfactoria o una convicción firme sobre la culpabilidad. Quizás sea algo así como “una confianza en la culpabilidad mayor del 90%.” Por su parte, en el sistema continental, los juegos de palabras correspondientes refieren cosas como una íntima convicción acerca de la culpa del acusado, o un sentido de certeza alcanzado libremente. Lo que éstas y otras variantes familiares comparten es el uso de la existencia de una alta confianza personal respecto de la historia de la acusación que funciona como piedra de toque y señal para decidir si la culpabilidad ha sido probada⁶. El Tribunal Supremo de los EEUU ha reconocido numerosas veces que la regla actual para decidir sobre la culpabilidad o inocencia es un criterio enteramente subjetivo (sin, añadiría yo, objetar el hecho de que lo es en demasía). En el caso *Jackson v. Virginia*, el Tribunal Supremo insistió que los jurados, para poder condenar a alguien, deben lograr “un estado subjetivo cercano a la certidumbre.”⁷. Fuera del derecho, los EdP nunca son formulados en términos de la confianza subjetiva de los investigadores sino en términos de los tipos de conexiones lógicas que deben existir entre la evidencia disponible y las hipótesis en cuestión a los efectos de considerar dicha hipótesis como probada. Fuera del derecho, la confianza racional en una conjetura sigue a su prueba, nunca la

⁶ Thomas Starkie, el gran teórico jurídico del siglo XIX, capturó el sentimiento generalizado que estoy criticando cuando insistió en que: Qué circunstancias se sumarán a la prueba nunca puede ser [un] el asunto de una definición general; el parámetro jurídico es la suficiencia de la evidencia para satisfacer y concienciar al jurado. STARKIE, Thomas *Un tratado práctico de la Ley de la evidencia* 513 (Benjamin Gerhard Ed., sexta edición, 1842) (1837).

⁷ *Jackson v. Virginia*, supra, 443 US at 315. Sentimiento similares pueden encontrarse inter alia en *Accord Johnson v. Louisiana*, 406 US 356 at 360, (1972) y *In re Winship*, 397 US at 364 (1970).

precede. Pero, dentro del derecho, tal confianza precede, certifica, e incluso constituye, la prueba.

Ahora bien, alguien podría cuestionarse ¿es que acaso la regla de la prueba más allá de una duda razonable no especifica un estándar de prueba genuino? La respuesta, como mostrará una pequeña reflexión, es un no rotundo. Como ya he defendido más extensamente en otra parte⁸, en la práctica actual de los Estados Unidos y de otros países del *common law* la duda razonable está completamente indefinida o definida de una manera tan imprecisa que resulta enteramente inútil. He mostrado que de hecho el estándar actual de la prueba en un juicio penal americano es simple y exclusivamente la convicción firme de la culpabilidad. El sistema no le ofrece al jurado EdP neutrales y objetivos y establece que es la intensidad de su confianza subjetiva en la culpabilidad la que determina si debe condenar o absolver al acusado. Para hacer las cosas aún peor, el sistema tampoco establece controles para saber cómo el jurado va a alcanzar ese nivel subjetivo de confianza. Éste confiere enteramente rienda suelta al jurado para hacer de la evidencia lo que se quiera y lo único que al final se reclama es que, si se vota por la condena, entonces se esté persuadido genuinamente de que el acusado cometió el crimen. Consecuentemente, si se absuelve al acusado, entonces cabría suponer que el jurado tiene una duda razonable acerca de su culpabilidad pero, dado que los tribunales hacen todo lo posible para insistir en que los jurados no necesitan ser capaces de articular sus dudas, las absoluciones quedan limitadas a un arbitrario y subjetivo conjunto de convicciones. Lo que observamos aquí no es un EdP sino una excusa o un pretexto débil para condenar o absolver.

Pero supongamos, para aclarar el tema, que el derecho penal tuviese un estándar genuino de prueba, que no dependiera de la evaluación subjetiva de la culpabilidad por parte de los jurados sino de que la acusación establezca una conexión deductiva sólida entre la evidencia presentada y la culpabilidad del acusado. En tales circunstancias, la culpabilidad no dependería de la introspección de los jurados, de su confianza en la culpa, sino de la determinación que ellos hagan acerca de si el EdP ha sido satisfecho.

Es justo preguntar: ¿qué significaría tener un estándar jurídico de prueba que no fuera parasitario de las tareas subjetivas de jurado respecto de su alta confianza ni de la alta probabilidad de culpabilidad? Para comenzar, recordemos el EdP que dominó el derecho Romano-Germánico de la tardía edad media a través del Siglo de las Luces. Dicho de manera simple, e ignorando los matices, la fórmula para asegurar una convicción en los juicios penales de aquellos tiempos fueron dos elementos: la presentación de dos

⁸ Véase mi artículo *Is Reasonable Doubt Reasonable?* 9, LEGAL THEORY 295 (2003).

testigos oculares del delito que fuesen creíbles o una confesión plausible del acusado. (Me adelanto aquí para agregar que no apruebo este estándar y que su uso en este apartado del trabajo tiene un carácter meramente ilustrativo.) Con este estándar en su lugar apropiado, la creencia subjetiva del juez (obviamente, los jurados no eran utilizados para juzgar los hechos) acerca de la inocencia o la culpabilidad del acusado carecía de relevancia. El EdP especificaba aquello que se necesitaba para condenar al acusado. Aunque este era un EdP poco confiable (en tanto que la ausencia de los dos testigos o de una confesión no era un signo seguro de la inocencia ni una confesión extraída mediante la tortura era un buen indicador de la culpabilidad), lo que debe decirse es que, por lo menos, era un EdP en el sentido estricto del término. La pregunta que tenemos que hacernos es si podemos idear algo similar alejado de lo que son nuestras prácticas contemporáneas.

Lo cierto es que ya han sido presentadas varias posibilidades interesantes que quizás puedan funcionar. Considere, por ejemplo, un EdP como el siguiente:

(A) Si es creíble la prueba acusatoria o un testimonio que resultaría difícil de explicar si el acusado fuese 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, tomando las indicaciones de algunas de las interesantes propuestas de Ron Allen⁹:

(B) 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.

No defiendo, para los propósitos presentes, los dos estándares propuestos. Simplemente los cito porque ambos exhiben una virtud clave de la que carece la práctica actual en los juicios penales; a saber, ellos le dicen al jurado lo que debe buscar en las pruebas para justificar la condena de un acusado. No piden que el jurado entre en un meta-análisis de su propio grado de confianza respecto de la culpabilidad o inocencia del acusado. (Puede ser, por supuesto, que en muchos de los casos en los que estarían satisfechos los estándares (A) o (B) la mayoría de los jurados concluyan también con una convicción firme acerca de la culpabilidad del acusado. De hecho uno esperaría que este fuera el caso. Pero esto no es necesario.) Curiosamente ni el estándar (A) ni el (B) representan una interrupción radical con respecto a las prácticas históricas del *common law*. Durante mucho tiempo, más con-

⁹ Véase especialmente ALLEN, Ron *Factual Ambiguity and a Theory of Evidence*, 88 Nw. L. Rv. 604 (1994) Y del mismo autor *Constitutional Adjudication, the Demands of Knowledge and Epistemological Modesty*, 88 Nw. U. L. Rev. 436 (1993).

cretamente hasta 1954 cuando el Tribunal Supremo lo eliminó de los juicios federales, había una instrucción oficial al jurado para los casos que implicaban pruebas circunstanciales. La línea de esta instrucción para los casos de pruebas circunstanciales todavía se utiliza en los tribunales del Estado de California. Dónde se lee, en parte:

Sin embargo, el hallazgo de la culpabilidad respecto de la comisión de un delito no puede basarse en pruebas circunstanciales a menos que las circunstancias probadas sean no sólo (1) coherentes con la teoría de que el acusado es culpable del crimen sino que además, (2) no puedan ser reconciliadas con cualquier otra conclusión racional. Si las pruebas circunstanciales permiten dos interpretaciones razonables, una que señala a la culpabilidad del acusado y la otra a su inocencia, usted deberá adoptar esta segunda interpretación y rechazar la interpretación que señala la culpabilidad¹⁰.

El lenguaje específico de esta regla es redundante (ya que la segunda oración está contenida en la primera) e inapropiado (en tanto que la condición (2) debería leerse algo como “son inconsistentes con la teoría de que el acusado es inocente del delito”). Pero tal desorden puede aclararse fácilmente. Lo que esta regla insinúa claramente es que (en los casos que implican pruebas circunstanciales) una convicción funciona bien sólo si las evidencias excluyen cualquier hipótesis razonable excepto la de la culpabilidad material del acusado. Esta instrucción tiene la virtud de ir directamente a la pregunta de la naturaleza de la prueba requerida y evita el devaneo de caracterizar cómo los jurados pueden sentirse acerca del asunto. Lo que les dice a los jurados es que no deben condenar a menos que la evidencia haya refutado previamente cualquier versión de los hechos que les permita pensar en la inocencia del acusado. Les dice:

(C) Resuelva si los hechos establecidos por la acusación refutan cualquier hipótesis aún ligeramente razonable que usted puede pensar respecto de la inocencia del acusado. Si ellos lo hacen, usted debe condenarlo. De otro modo, usted debe absolver.

Cabe resaltar que los tres EdP instan al juzgador de los hechos a probar rigurosamente (para utilizar una frase popperiana familiar) la hipótesis de la culpabilidad antes de que la adopten. Contamos con una prueba rigurosa de una hipótesis, H, cuando tenemos una evidencia de que sería muy sorprendente o improbable que H fuese falsa. Lo que los jurados deberían buscar no es la prueba que meramente confirma o es compatible con la hipótesis de la culpabilidad sino la evidencia de que no se espera que la hipótesis sea falsa. Los EdP (A)–(C) señalan precisamente en esa dirección.

¹⁰ CALJIC 2.01.

Debemos notar que (C) no pide que los jurados hagan una lectura de su confianza acerca de la culpabilidad más de lo que (A) o (B) lo hacen. Sino que les pregunta si la evidencia creíble presentada en el juicio excluye cada hipótesis *prima facie* plausible de la inocencia que puedan imaginar. La propuesta de los estándares de Allen, (B) y (A), que implican el elemento de la prueba de lo inesperado, está claramente dentro de la misma familia. ¡Pero, en vez de continuar con el uso de la instrucción para las pruebas circunstanciales, el Tribunal Supremo de los Estados Unidos decidió, de forma increíble, que decirles a los jurados que podían condenar sólo si estaban firmemente convencidos de la culpabilidad del acusado era la instrucción correcta y que el estándar de pruebas circunstanciales estaba equivocado!¹¹ Aunque esta decisión fue coherente con la creencia de este Tribunal de que la certeza subjetiva es una condición *sine qua non* para una sentencia de culpabilidad, supuso, sin embargo, un paso hacia atrás que oscureció aún más un proceso ya de por sí oscuro.

Obviamente, ninguno de los tres estándares aquí discutidos depende de la estimación subjetiva del jurado respecto de la probabilidad de culpabilidad del acusado. De hecho, cabría una situación en la que una jurado está obligada a condenar, incluso si ella (quizás de forma aviesa) sigue creyendo en la inocencia del acusado a pesar de que el estándar ha sido satisfecho. Quizás podría pensarse que los jurados serían reticentes ante la sugerencia de verse obligados a condenar a alguien que ellos consideran inocente. Pero este dilema no es de un tipo diferente a los de la situación actual en la que los jurados a menudo se encuentran obligados a declarar inocente a alguien a quien ellos creen culpable. Si los jurados pueden comprender (como obviamente hacen) que el requerimiento de la prueba debe triunfar sobre sus creencias privadas en lo que concierne a la culpabilidad, también pueden aprender seguramente que sus creencias privadas acerca de la inocencia tendrán que ceder igualmente ante un EdP plausible.

Enfrentados con EdP como los propuestos tales como (A), (B), o (C), la reacción inicial es sostener que las cosas posiblemente no pueden ser ni tan sencillas ni tan ‘formularizables’ como ellos sugieren. Tratamos, a fin de cuentas, con acontecimientos humanos sumamente complejos y, en el caso de la deliberación del jurado, con procesos de pensamiento que no son claros. Una vez más, de forma constructiva, el instinto deberá tratar de imaginar los casos donde (A) a través de (C) fallaría al condenar a un culpable o, peor, donde quizás condene a un inocente. Tales casos no son tan difíciles de imaginar. Pero esto está más allá del punto en cuestión. Fuera

¹¹ *Holland v. US*, 348 US 121(1954). La Corte opinó que la instrucción para las pruebas circunstanciales era confusa e incorrecta. *Id.*, at 139-140.

de las matemáticas y de la lógica formal, no tratamos con la certeza ni con relaciones deductivas entre premisas y conclusiones. En vez de eso, encaramos la incertidumbre y las inferencias falibles. Es parte de la naturaleza ampliativa de las reglas de inferencia la posibilidad de conducir a conclusiones erróneas. *Ningún* EdP puede eliminar completamente la incertidumbre ni superar la clase de inferencias abductivas utilizadas en el razonamiento legal preservador-de-la-verdad.

La pregunta pertinente concierne a la *frecuencia relativa* de los errores que los estándares (A)-(C) producirían en comparación al actual mecanismo que hace depender el veredicto de un juicio del instinto desordenado e indisciplinado del jurado acerca de la culpabilidad e inocencia o bien que lo hace depender de una arbitraria atribución de probabilidades por parte de los jurados. Desgraciadamente, nadie ha reunido la información empírica oportuna para establecer esta pregunta de una manera definitiva. Pero *podría* lograrse mediante una investigación empírica. Mientras tanto, hay algunas razones plausibles para creer que reglas como (A)-(C) funcionarían mejor que cualquier criterio subjetivo. Considere, por ejemplo la regla (A). Lo que ella pregunta es si la acusación ha logrado presentar un testimonio creíble y pertinente o una evidencia física que serían muy difíciles de explicar si el acusado fuese inocente.

Es una regla general del razonamiento común acerca del mundo que *cualquier* hipótesis histórica se prueba mejor si los hechos sobre aquello que tendría que explicar pero no puede son relevantes y están claramente establecidos. Esta estrategia de la evaluación de la hipótesis ocurre constantemente en las llamadas ciencias históricas, como la historia de la humanidad o la historia de la tierra. Si, por ejemplo, queremos evaluar la hipótesis según la cual la colisión de un meteorito fue la causa de la extinción de los dinosaurios, nos preguntaremos si hay alguna evidencia creíble datada en el período pertinente que sería enteramente explicable por la hipótesis del choque pero que resulte ininteligible si (por ejemplo) los dinosaurios desaparecieron debido a la pobre ventilación de sus cuerpos. Igualmente nos cuestionaremos si hay hechos bien-fundados acerca de la extinción de los dinosaurios que no ponen en cuestión la hipótesis de la colisión. Desde que hay tal evidencia (la famosa prueba del "límite de iridio") explicable por la hipótesis de la colisión, y ningún hecho notable no explicable por ella, tenemos razones poderosas para acreditar la hipótesis del meteorito.

Este tipo de indagación no es diferente de lo que pasa diariamente en los tribunales de justicia. La hipótesis sobre la culpa y la inocencia se descartan si uno mira hacia la evidencia para encontrar pruebas probatorias de esas hipótesis. Lo que un EdP como (A) hace es explicitar de manera completa la clase de prueba que los jurados deben buscar si cabe esperar una convic-

ción. En resumen, la clase de inferencia que invita a hacer el EdP (A) no es nueva ni extraña en una corte de justicia, ni sería distinta a lo que los jurados hacen en su vida cotidiana. Lo que hasta ahora los tribunales no han hecho es decir que, si una regla como (A) ha sido satisfecha por el caso de la acusación, entonces el jurado tiene una prueba de la culpabilidad y si no ha sido satisfecha, entonces no tienen una prueba. ¡Al contrario, varios tribunales de apelación han dicho explícitamente que los estándares como los hasta aquí considerados son una exigencia excesiva! Por ejemplo, la Primera Corte Federal de Apelación de los EEUU ha defendido que: “La acusación puede demostrar su caso mediante pruebas circunstanciales, y no necesita excluir cada hipótesis razonable de inocencia siempre que la evidencia total permita una conclusión de culpabilidad más allá de un duda razonable.”¹². Pero, uno de los misterios que rodean al BARD es cómo imagina el Tribunal que un jurado puede estar seguro más allá de una duda razonable de la culpabilidad del acusado aun cuando igualmente cree que hay “una hipótesis razonable de inocencia” que no ha sido excluida. Ocho años antes ese mismo Tribunal había argumentado que: “El juzgador del hecho es libre de escoger entre varias construcciones razonables de la evidencia.”¹³. ¿Cómo puede un jurado estar seguro más allá de una duda razonable de la culpabilidad del acusado, cuándo hay “varias construcciones razonables de la evidencia,” algunas de las cuales son exculporias? Si el BARD es tan permisivo como esto, entonces es urgente su reemplazo por un EdP más exigente como (A)-(C).

Las reglas como (A)-(C), si son adoptadas como EdP, resolverían los dos problemas que aquí hemos señalado. Estos son genuinos EdP no subjetivos y, a diferencia tanto del BARD como de un valor concreto de la probabilidad, no dejan mucho margen de ambigüedad respecto de si han sido satisfechos. Y lo que es aún más importante, requieren que, antes que los jurados puedan condenar a un acusado, se debe haber sometido la hipótesis de la culpabilidad a una prueba rigurosa. Deben haber mostrado que la hipótesis de la inocencia simplemente no puede ser encajada con la evidencia presentada en el juicio.

La relación de un EdP riguroso con otros principios de la distribución del Error

Dije al principio que el EdP era uno de varios conceptos familiares que procuran distribuir los errores en favor del acusado. Los otros conceptos obvios son la presunción de inocencia, la carga de la prueba de la acusación, y el beneficio de la duda. Por todo lo hasta aquí dicho, creo que si relacioná-

¹² US v. Gabriner, 571 F.2d 48, 50 (1st Cir., 1978).

¹³ US v. Thornley, 707 F.2d 622, at 625 (1st Cir., 1983).

semos un EdP con dichos conceptos esto tendría repercusiones significativas respecto de cómo los entendemos.

Considere, en primer lugar, el beneficio de la duda. A menudo se dice que el acusado tiene siempre el derecho del beneficio de la duda. Esta es una manera no atinada de decirlo, *ya que siempre hay alguna duda residual*. Si debiéramos dar a un acusado todo el beneficio de la duda, sería imposible condenar a alguien. Por esta razón, tenemos que decidir a cuánto beneficio de la duda tiene derecho. Responderé a esta pregunta de una manera directa: un EdP alto ya incorpora un beneficio de la duda inmenso. Éste le indica al acusado: “Usted será absuelto, incluso si pensamos que es probablemente culpable, a menos que el nivel de su culpabilidad satisfaga un EdP muy exigente.”. En resumen, no hay beneficio de la duda que sea de fundamentación-libre, independientemente del EdP. El nivel apropiado del beneficio de la duda está ya incorporado en un EdP favorable-al-acusado. Así, dado un estándar exigente, el principio de *in dubio pro reo* habrá sido ya incorporado en el juicio y no tendrá un rol diferenciado más allá del propio EdP.

Considere, ahora, la presunción de inocencia. Los tribunales, especialmente en el mundo anglosajón, han tendido a suponer que la presunción de inocencia requiere que el juzgador de los hechos comience el juicio creyendo que el acusado no cometió el crimen, esto es, que es *materialmente* inocente. Con un EdP apropiado, la presunción de inocencia material no sería más necesaria ni apropiada. Lo que sería exigible es que el juzgador de los hechos crea que la culpabilidad del acusado tiene que ser probada (lo que llamo la presunción de inocencia *probatoria*), y que la única prueba relevante para su culpabilidad o inocencia es la prueba que será producida en el juicio¹⁴.

Finalmente, ¿qué sucede con el otro instrumento de distribución del error, la carga de la prueba de la acusación? Pues que se convierte en enteramente superfluo. Un EdP exigente ya deja claro que la culpabilidad del acusado debe ser probada en un nivel muy elevado y que el acusado debe ser absuelto, si no se satisface el EdP. En tanto que podemos suponer, que el acusado no ofrecerá pruebas sólidas de su propia culpa, la carga de demostrar la culpabilidad conforme al estándar pertinente inevitablemente caerá en el acusador.

En resumen, básicamente el único mecanismo para la distribución de los errores que se requiere en un juicio es un estándar no-subjetivo de la prueba. Si ponemos ese estándar a la altura correcta, capturará completamente nuestras creencias compartidas acerca de la ratio apropiada de absoluciones

¹⁴ Desarrollo este argumento con mayor detalle en un trabajo titulado *La Presunción de la Inocencia: material o probatoria?*, que será publicado en la revista *Legal Theory*.

falsas por falsas convicciones. Cualquier otra cosa que maquille los elementos epistemológicos de un juicio penal (específicamente, las reglas de la evidencia y del procedimiento) puede entonces ser escogida a la luz del núcleo duro con la intención de aminorar la probabilidad de juicios erróneos antes que con la perspectiva de lograr distribuirlo de una forma favorable-al-acusado.

Confío que esta conclusión le golpeará de una manera tan radical como creo que es. Para ser más explícito, esto supone que las reglas de la evidencia y del procedimiento deberían ser diseñar con miras a aminorar el error judicial y no con el objetivo de reducir la posibilidad de una absolución falsa (que es una cuestión que concierne exclusivamente al tema de la distribución). Las reglas existentes en todo sistema jurídico, al menos de los que conozco, están llenas de esfuerzos, bien intencionados pero mal entendidos, encaminados a rebajar la escala de justicia con el fin de que sea más difícil obtener la culpabilidad.

(Trad. de Raúl Calvo Soler)



DOXA 28 (2005)
