

ENCUENTROS CERCANOS CON ARGUMENTOS DEL “TERCER TIPO”: RAZONAMIENTO PLAUSIBLE Y PROBABILIDAD SUBJETIVA COMO MODELOS DE EVALUACIÓN DE ARGUMENTOS

CHRISTIAN GAVIRIA

Departamento de Psicología

Universidad de los Andes

cgaviria@uniandes.edu.co

WILLIAM JIMÉNEZ-LEAL

Departamento de Psicología

Universidad de los Andes

w.jimenezleal@uniandes.edu.co

RESUMEN: Este artículo presenta un análisis comparativo de los modelos de argumentación basados en las nociones de probabilidad subjetiva y de razonamiento plausible. Se hacen explícitos los “parecidos de familia” entre probabilidad y plausibilidad, y se examinan las diferencias en las prescripciones invocadas para la evaluación de tres tipos de falacias informales: apelación a la autoridad (*ad verecundiam*), a la popularidad (*ad populum*) y petición de principio (*petitio principii*). Se concluye que el razonamiento plausible, como Rescher y Walton lo describen, no proporciona una alternativa sólida a la probabilidad como modelo normativo o descriptivo de la evaluación de argumentos.

PALABRAS CLAVE: lógica informal, falacias, Bayes, argumentación, ciencia cognitiva

SUMMARY: This paper presents a comparative analysis of argumentation models based on the concepts of subjective probability and plausible reasoning. This analysis makes explicit the “family resemblance” between subjective probability and plausible reasoning, while examining the differences in the requirements that each model invokes regarding the evaluation of three types of informal fallacies: argument from authority (*ad verecundiam*), appeal to popularity (*ad populum*) and begging the question (*petitio principii*). We conclude that plausible reasoning, as it is characterized by Rescher and Walton, does not provide a strong alternative to probability as either a normative or descriptive model of argument evaluation.

KEY WORDS: informal logic, fallacies, Bayes, argumentation, cognitive science

1. *Introducción*

El adjetivo “plausible” se usa cotidianamente como calificativo de afirmaciones, argumentos y decisiones. Sin embargo, dicho término connota significados distintos dependiendo del contexto. Así, cuando en el marco de un proceso legal se sostiene que es plausible que cierto delito haya tenido un móvil pasional, no se está afirmando categóricamente que éste sea efectivamente el móvil; más bien se alude a que esta proposición se acepta como punto de partida de

la investigación, hasta que la evidencia presentada por alguna de las partes haga necesario reconsiderarla. Decir que una teoría científica es plausible significa que es coherente con el conocimiento acumulado en el dominio pertinente al fenómeno que se intenta explicar, con independencia de si existe, o no, evidencia directa que confirme dicha teoría. Un curso de acción es plausible si los agentes que participan pueden convenir en que es la alternativa más razonable para lograr sus fines a partir de los recursos disponibles, aunque no exista certeza de tal logro. ¿Qué tienen en común estos distintos usos? ¿Qué significa decir que un enunciado, o un argumento, es plausible? ¿Qué relación existe entre lo “plausible”, lo “probable” y lo “verdadero”? Si bien no existe consenso en la literatura filosófica que permita dar una respuesta definitiva a estas preguntas,¹ es posible acotar los límites del debate en torno a ciertos aspectos característicos. Con esto se persigue un doble objetivo: primero, precisar la noción de plausibilidad; segundo, contrastar la forma en que las nociones de plausibilidad y probabilidad han dado origen a programas empíricos de investigación en ciencia cognitiva sobre argumentación. Estos objetivos no son independientes, y se pone aquí la elucidación del concepto de plausibilidad al servicio de la fundamentación de una teoría psicológica adecuada sobre argumentación.

El presente documento está organizado en cinco secciones. En la primera, se examinarán cinco propiedades típicas del concepto de plausibilidad, las cuales tienen raíces históricas profundas y se mencionan de manera recurrente en la literatura filosófica: 1) La relatividad a una audiencia; 2) la representación de la fuerza argumentativa en una escala numérica; 3) la sensibilidad a la introducción de nueva información; 4) la posibilidad de realizar inferencias a partir de contradicciones, y 5) la separación entre lo plausible y lo verdadero. En segundo lugar, se abordará la cuestión de hasta qué punto estas propiedades permiten distinguir claramente los conceptos de plausibilidad y probabilidad subjetiva. En tercera instancia, se presentarán los modelos de evaluación de argumentos basados en las nociones de probabilidad subjetiva y de plausibilidad, haciendo énfasis en sus diferencias normativas. En la cuarta parte, se describirán tres escenarios argumentativos con respecto a los cuales los dos modelos en cuestión hacen diferentes prescripciones a propósito del carácter falaz o no falaz de los argumentos, mostrando cómo la aplicación de las reglas del razonamiento plausible puede conducir a resultados contrarios a la intuición. Por último, a manera de conclusión, se

¹ Véase Blair 2011.

exponen hallazgos empíricos recientes que sugieren que la manera en la que las personas efectivamente evalúan argumentos informales parece ajustarse mejor al modelo basado en probabilidades subjetivas que al razonamiento plausible.

2. Lo plausible, lo probable y lo verdadero

Relatividad a una audiencia. Etimológicamente, el vocablo “plausible” se deriva del latín *plausibilis*, que significa “digno de aplauso”. El Diccionario de la Real Academia Española (2001) incluye las acepciones “atendible”, “admisible” y “recomendable” en su definición de “plausible”. Este uso pone ya de manifiesto una de las propiedades características del concepto: una afirmación siempre es plausible o implausible para alguien. En efecto, la plausibilidad no es una propiedad semántica de una afirmación o argumento, sino que describe una relación pragmática entre un enunciado o argumento y una audiencia determinada (Vega-Reñón 1998). De acuerdo con Walton (2008), esta idea tiene sus orígenes en los sofistas de la Grecia antigua, quienes usaban el vocablo *εἰχός* para designar argumentos cuya aceptación se basa en el conocimiento compartido por el proponente y la audiencia acerca de lo que sería “normal” o “razonable” esperar en una situación familiar para ambas partes. Así, en principio, puede definirse lo plausible como aquello que resulta “común”, “familiar”, “esperable” o “razonablemente presumible” para una audiencia determinada. En el mismo sentido, Aristóteles emplea el término *ἔνδοξα* para referirse a las proposiciones de las que “se sabe que la mayoría de las veces ocurre así o no ocurre así, o es o no es” (*A.Pr.* 70a3–5), y a aquellas “que les parecen bien a todos, o a la mayoría, o a los sabios, y, entre estos últimos, a todos o a la mayoría, o a los más conocidos y reputados”. (*Top.* 100b21–23). En este orden de ideas, se pueden distinguir dos tipos de plausibilidad, dependiendo de si se consideran argumentos (*εἰχός*) o proposiciones (*ἔνδοξα*).

Siguiendo el análisis de Vega-Reñón (1998), el orden en que Aristóteles enumera los grupos de personas para los cuales una proposición es aceptable, no es fortuito: se trata de una escala gradual decreciente que opera bajo el supuesto de que una afirmación es más plausible en tanto que más personas la acepten como tal. Así, de la relatividad a la audiencia de los juicios de plausibilidad y la necesidad de hacerlos comparables entre sí surge la representación de los juicios de plausibilidad como valores en una escala ordinal. Como se verá más adelante, la formalización del razonamiento plausible desarrollada por Nicholas Rescher (1976) intenta capturar esta intuición al

asignar valores ordinales de plausibilidad a proposiciones y conjuntos de proposiciones.

En el libro segundo de la *Retórica*, Aristóteles proporciona un ejemplo de argumento “eikótico”, el cual es recogido por Walton (2008): supóngase un juicio en el que se intenta definir quién lanzó el primer golpe en una riña entre dos hombres. Uno de ellos es visiblemente más pequeño y débil que el otro. El alegato del primero consiste en cuestionar la hipótesis de que un hombre como él, con plena conciencia de sus actos, inicie deliberadamente una pelea sabiendo que no tiene oportunidad alguna de ganarle a su contendor y que probablemente saldrá lastimado. Con base en este argumento y en ausencia de evidencias adicionales, el jurado resuelve que es poco plausible que el hombre más débil sea quien ha iniciado la riña. Este ejemplo ilustra una implicación importante de la relatividad de la plausibilidad a una audiencia: supóngase que un juicio idéntico al anterior se lleva a cabo en una comunidad en la que se castiga duramente a los hombres fuertes que golpean a los más débiles, pero no a la inversa. En este caso, el argumento del hombre más débil resulta menos plausible, dado que sus oportunidades de ganar la pelea aumentarían en la medida en que el hombre fuerte puede verse más perjudicado si responde a sus provocaciones. Así, aquello que resulta plausible o aceptable para una audiencia determinada en un marco compartido de conocimientos, presunciones, valores, etc., puede no serlo para otra que difiera en aspectos relevantes de dicho marco. Lo que resulta plausible para un panel de expertos en un tema determinado (*e.g.*, la dualidad onda-partícula de la luz o la existencia de una partícula subatómica que transmita la fuerza gravitatoria) puede ser poco plausible para una audiencia de no expertos (Wolpert 1994). Estos ejemplos ponen de manifiesto la importancia de la base de conocimiento relevante y disponible para la audiencia como un factor decisivo en sus juicios de plausibilidad. Como se verá más adelante, dicho factor está relacionado con otro rasgo distintivo del concepto: la sensibilidad a la introducción de nueva información.²

² Un elemento afín lo constituye la relación entre “plausible” y “persuasivo”, en la medida en que ambos aspectos son juzgados por una audiencia. Hay al menos una diferencia entre ambas nociones: un argumento es persuasivo en tanto que logre generar un cambio significativo en las actitudes y/o creencias de la audiencia con respecto a un objeto determinado, mientras que el fin último de un argumento plausible es algo más modesto: lograr que el interlocutor acepte de manera tentativa un conjunto de proposiciones en calidad de presunciones razonables (Vega-Reñón 1998; Walton 2008). No obstante, es evidente que la relación entre plausibilidad y persuasión aún no ha sido analizada lo suficiente.

Revocabilidad de las inferencias. Mientras que la plausibilidad de una proposición o argumento puede ser evaluada de manera distinta por distintas audiencias, también es posible que una afirmación resulte plausible para una misma audiencia en un momento determinado e implausible en otro momento, o viceversa. Para que esto ocurra, tiene que producirse algún cambio en la base de conocimiento compartido por los participantes del proceso argumentativo. En la literatura filosófica, la posibilidad de que inferencias previas sean modificadas en virtud de la introducción de nueva información se denomina comúnmente “revocabilidad” (*defeasibility*) o no monotonicidad. Pollock (1995) destaca esta característica como aquella que distingue el razonamiento plausible del deductivo. De acuerdo con Pollock, un argumento es revocable cuando sus premisas son razonables en principio (*prima facie*), pero es posible que la introducción de premisas adicionales, denominadas “revocadores” (*defeaters*), ocasione que las premisas originales sean verdaderas y la conclusión falsa (Pollock 1992). En este contexto, un “revocador” puede ser o bien un argumento alternativo cuya conclusión sea la negación de la conclusión del argumento a revocar, o bien una proposición que implique la negación del nexo inferencial que existe entre las premisas y la conclusión del primer argumento. Para ilustrar la diferencia entre una inferencia deductiva y una revocable, Pollock (1995) propone el siguiente ejemplo: supóngase una habitación oscura en la que hay un objeto iluminado por una luz roja. Si el objeto se ve rojo, un observador desprevenido podría realizar la siguiente inferencia:

Cuando un objeto luce rojo, entonces es de color rojo.

El objeto *X* luce rojo.

Por lo tanto, el objeto *X* es de color rojo

No obstante, un escéptico podría formular el siguiente contraargumento:

Cuando un objeto es iluminado por una luz roja, el objeto puede lucir rojo aunque no lo sea.

El objeto *X* está iluminado por una luz roja,

Por lo tanto, el objeto *X* puede no ser de color rojo.

El segundo argumento no refutaría en sentido estricto la conclusión del primero, puesto que, si el objeto es rojo, luciría como tal al

ser iluminado por la luz roja. En lugar de ello, pone en cuestión la validez deductiva de la inferencia “Si el objeto x luce de color rojo, entonces el objeto x es de color rojo”, aunque no impide que dicha inferencia siga siendo considerada plausible. Walton (2008) define la forma básica de un condicional plausible de la siguiente manera: “Si A es verdadero y se mantienen otras cosas iguales, entonces se puede esperar normalmente que B sea verdadero también, sujeto a excepciones” (p. 97).³ La expresión “sujeto a excepciones” permite que en algunos casos el antecedente sea verdadero y el consecuente falso sin que esto implique la falsedad de la inferencia plausible. No obstante, como se verá a continuación, los condicionales plausibles siempre son susceptibles de ser desvirtuados como presunción razonable a partir de la introducción de nueva evidencia.

La revocabilidad, o no monotonicidad, de las inferencias plausibles permite que la introducción de nuevos elementos a un conjunto determinado de proposiciones en un momento dado cambie la plausibilidad de las proposiciones que ya formaban parte de dicho conjunto. Esta propiedad responde al hecho de que la evidencia en la que se basan la mayor parte de las inferencias cotidianas es usualmente incompleta y algunas veces resulta poco informativa e incluso contradictoria. En este sentido, el concepto de plausibilidad recoge un aspecto esencial de contextos argumentativos contenciosos, tales como los litigios judiciales, en los cuales cada una de las partes recopila y presenta nuevas pruebas en diferentes fases del juicio, con el fin de sustentar sus propios argumentos y cuestionar los de la contraparte. A su vez, este proceso obliga a quien juzga (juez o jurado) a reconsiderar permanentemente la plausibilidad de los argumentos de ambas partes en función de los cambios en el acervo probatorio. En el caso del argumento eikótico antes mencionado, si la defensa del hombre corpulento presenta pruebas de que el hombre más pequeño es un boxeador entrenado y ha tenido antecedentes penales por riña en vía pública y lesiones personales, el jurado debería reconsiderar la plausibilidad del argumento de este último a la luz de la nueva evidencia disponible (Walton 2006).

Inferencia a partir de contradicciones. La sensibilidad del razonamiento plausible a la introducción de nueva información trae consigo otra característica que la distingue de la lógica deductiva y de la inferencia inductiva: la posibilidad de realizar inferencias a partir de reportes contradictorios de información. Rescher (1976) plantea que una de las metas del razonamiento plausible es operar con conjuntos

³ Las traducciones de todas las citas de este artículo son nuestras.

inconsistentes de proposiciones, intratables desde la lógica deductiva y el cálculo de probabilidades. Así, si se tienen dos argumentos internamente coherentes, uno de los cuales implica la proposición *A* y otro la proposición *No A*, es posible que los índices de plausibilidad de las premisas de cada argumento sean suficientemente altos o bajos como para que *A* y *No A* sean simultáneamente plausibles o implausibles. Por ejemplo, en la situación descrita en el argumento eikótico, son presentados dos argumentos de plausibilidad similar y conclusiones contradictorias (e.g., el hombre fuerte/el hombre débil inició la riña), los cuales provienen de dos fuentes con intereses claramente opuestos.⁴

En las inferencias probabilísticas se asignan valores distribuyendo una cantidad fija (exactamente una unidad) entre un conjunto exhaustivo de proposiciones mutuamente excluyentes, en función del contenido semántico de cada proposición. En contraste, la asignación de valores de plausibilidad depende únicamente de la relación que existe entre la audiencia y la fuente o base de conocimiento de la cual proviene cada proposición o conjunto de proposiciones (Walton 2002; Vega-Reñón 1998). Una proposición podría ser plausible con respecto a una determinada base de conocimiento o fuente mientras que su negación podría ser igualmente plausible con respecto a otra base o fuente. Por tanto, desde la lógica de la plausibilidad, no es un sinsentido aceptar simultáneamente dos argumentos contradictorios de manera tentativa en un momento dado, teniendo en cuenta que la información disponible puede no ser suficiente para descartar alguno de los dos y que cualquiera de ellos es revocable en función de la introducción de nueva información (Walton, Tindale y Gordon 2014).

Plausibilidad y verdad. El último rasgo característico del concepto de plausibilidad puede encontrarse ya en la distinción aristotélica entre *εικός* y *ἀληθεία*, esto es, entre lo plausible y lo verdadero. Así, al criticar a algunos filósofos de su tiempo, Aristóteles señala que “hablan con verosimilitud (*εικαστος*), pero no hablan con verdad (*ἀληθή*).” (*Met.* 1010a5–6). En efecto, la caracterización de la inferencia plausible como la aceptación racional, tentativa y revocable de “lo que parece ser cierto” para una audiencia determinada, junto con la posibilidad de que una regla condicional plausible conduzca a una conclusión falsa, se aparta de la idea de fuerza inductiva y de

⁴ Nótese que la contradicción entre argumentos eikóticos es posible para Aristóteles, pero no así la contradicción entre proposiciones *plausibles*, fuente de problemas dialécticos (e.g., *Top.* 104a14; *SE* 177a12–14, 19–26). De aquí en adelante nos centramos en la plausibilidad de argumentos, y argumentos plausibles e incompatibles. Agradecemos a un árbitro anónimo hacernos notar este importante punto.

la noción de “transmisión de la verdad” implicada en el concepto de validez deductiva. Un argumento es deductivamente válido si es imposible que sus premisas sean verdaderas y su conclusión falsa, mientras que un argumento es inductivamente más fuerte mientras la probabilidad de que sus premisas sean verdaderas y su conclusión falsa sea menor. En contraste, en un argumento plausible las premisas no proveen evidencia para la conclusión en el mismo sentido que en un argumento inductivo probabilístico. Las inferencias plausibles no son evaluadas en términos de su capacidad para transmitir la verdad o aproximarse a lo verdadero, sino en virtud de su coherencia con un conjunto determinado de proposiciones o creencias (Woods y Gabbay 2005). En palabras de uno de los primeros comentaristas de Aristóteles, Alejandro de Afrodisia (*In Top.* 19.22–27; citado por Vega-Reñón 1998):

La plausibilidad difiere de la verdad no por ser falsa —algunas opiniones plausibles son de hecho verdaderas— sino por el criterio en el cual se basa el juicio. En el caso de la verdad, el juicio se hace con referencia a la cosa misma de la cual trata la opinión: cuando las opiniones coinciden, esto es verdadero. En el caso de la plausibilidad, el juicio se hace no con referencia a las cosas mismas, sino con referencia a los oyentes y sus presunciones sobre las cosas. (p. 104)

En este sentido, es posible que una proposición altamente plausible en un momento determinado resulte a la larga ser falsa y que una muy implausible resulte ser verdadera.

Resumen: plausibilidad y probabilidad subjetiva. Recientemente ha surgido un debate en la literatura sobre lógica informal y teoría de la argumentación en torno al estatus de la plausibilidad como un “tercer” tipo de proceso inferencial y estándar evaluativo, discernible de la lógica deductiva y el cálculo probabilístico. En el contexto de esta discusión, la plausibilidad cumple un papel doble: como principio que permite la caracterización adecuada de la argumentación en general, y como elemento clave en las descripciones de los procesos de evaluación de argumentos que llevan a cabo las personas. Esto es, la plausibilidad tiene un carácter tanto normativo (permite establecer las pautas de aquello que se considera un argumento correcto) como descriptivo (suministra el lenguaje necesario para descripciones relevantes de los procesos argumentativos). En el escenario ideal, un modelo normativo adecuado de los procesos argumentativos debería proporcionar estándares de racionalidad con respecto a los cuales sea

posible contrastar y entender por qué las personas evalúan argumentos de la forma en que lo hacen (Hahn y Oaksford 2006). La comparación entre la aproximación probabilista y plausibilista se juega tanto en el terreno de la caracterización de los argumentos, como en el de la psicología de la argumentación, entendida esta como el conjunto de descripciones y explicaciones de la manera en que las personas usan diversos esquemas argumentativos. El examen de las diferencias y coincidencias entre los modelos considerados en este texto se realiza en el contexto de esta distinción.

De acuerdo con los defensores de las aproximaciones probabilísticas a la inferencia humana, la idea de plausibilidad puede ser completamente asimilable a la de probabilidad subjetiva y no constituye una forma alternativa de inferencia (Hahn, Oaksford y Harris 2013; Hahn y Oaksford 2006). Por el otro lado, si bien los teóricos de la plausibilidad han demostrado claramente las diferencias entre este concepto y la deducción, principalmente en términos de su no monotonicidad y su renuncia a la idea de validez (Pollock 2006; Rescher 1976; Walton 2008), la distinción entre plausibilidad y probabilidad subjetiva es más difusa. Mientras que la lógica deductiva no comparte con el concepto de plausibilidad ninguna de las cinco propiedades mencionadas anteriormente, el razonamiento plausible comparte varias características con la inferencia probabilística: ambos son sensibles a la audiencia, a la información previa, a la incorporación de nueva evidencia y la posibilidad de cuantificar la fuerza de los argumentos. No obstante, de acuerdo con Rescher y Walton, la inferencia plausible es especialmente relevante en situaciones de conflicto de opiniones (es decir, en presencia de información contradictoria) y ausencia de evidencia concluyente a favor de alguna de ellas. En la siguiente sección se presentarán los fundamentos conceptuales del razonamiento plausible y del modelo de evaluación de argumentos basado en la idea de probabilidad subjetiva, como un paso previo a la comparación de sus prescripciones a propósito de tres escenarios argumentativos específicos.

3. Plausibilidad y probabilidad como modelos de argumentación

El modelo bayesiano de evaluación de argumentos. El principio básico del modelo probabilístico de evaluación de argumentos es que éstos pueden ser entendidos como arreglos de premisas y conclusiones que tienen probabilidades subjetivas asociadas, las cuales expresan el grado de creencia de un agente en cada una de ellas en un momento dado (Hahn y Oaksford 2006; 2007). La tesis fuerte de este modelo es

que el proceso de actualización del grado de creencia en la conclusión de un argumento en función del grado de creencia en sus premisas se ajusta a lo prescrito por el teorema de Bayes:

$$P(H | E) = \frac{P(E | H)P(H)}{P(E | H)P(H) + P(E | \neg H)P(\neg H)}$$

Donde la fuerza percibida del argumento $P(H | E)$ ⁵ es una función del grado de creencia previa en la conclusión $P(H)$ y las probabilidades de que las premisas sean verdaderas tanto en caso de que la conclusión también lo sea, $P(E | H)$, como en caso de que no, $P(E | \neg H)$. La división entre estas dos últimas probabilidades, denominada razón de verosimilitud (*likelihood ratio*), proporciona un índice de qué tan informativa o diagnóstica resulta la evidencia en relación con la conclusión (Hahn y Oaksford 2012). En este contexto, un argumento se considera *fuerte* en la medida en que la evidencia presentada en las premisas sea más informativa con respecto a la verdad de la conclusión, $P(E | H) > P(E | \neg H)$ y *débil* en caso de que no lo sea, $P(E | H) = P(E | \neg H)$. A su vez, el teorema de Bayes se deriva de la definición clásica de probabilidad condicional, $P(A | B) = P(A \cap B)/P(B)$ siempre que $P(B) > 0$, y de los tres axiomas del cálculo de probabilidades estándar.

Ventajas descriptivas del modelo bayesiano de evaluación de argumentos. Desde el punto de vista psicológico, son varias las características de la aproximación bayesiana que la convierten en un candidato viable a modelo *descriptivo* de evaluación de argumentos (Hahn *et al.* 2013; Hahn y Oaksford 2006). En primer lugar, la idea de probabilidad como grado de creencia subjetiva hace posible dar cuenta de argumentos que involucran eventos singulares y/o hipotéticos, a los cuales hacen referencia gran parte de los argumentos cotidianos. En segundo lugar, el concepto de actualización bayesiana distingue entre la probabilidad que el agente asigna a la conclusión de manera previa a la evidencia presentada en el argumento (*a priori*) y la probabilidad de la conclusión una vez conocida la evidencia (*a posteriori*), lo cual permitiría dar cuenta de la actualización de creencias en función de argumentos y explicar por qué un mismo argumento puede ser evaluado de manera diferente por distintas audiencias. En tercer lugar, los modelos bayesianos proporcionan parámetros específicos que representan diferentes aspectos de la calidad y cantidad del *contenido* de las premisas y conclusiones de los argumentos, permitiendo entre

⁵ En el caso de la argumentación, H representa la conclusión y E el conjunto de premisas del argumento.

otras cosas identificar aquellos contextos en los cuales la escasez de información disminuye la fuerza de los argumentos y no produce efecto alguno sobre las creencias de los agentes. Desde un punto de vista más general, la parametrización del modelo hace posible derivar predicciones tanto *ordinales* (e.g., el argumento *X* tendrá una fuerza percibida superior/inferior a la del argumento *Y*) como *puntuales* (e.g., dado el conjunto de probabilidades previas asignadas, el grado de creencia posterior en la conclusión *C* a la luz de las premisas *P₁* y *P₂* será exactamente *x*). Esta ventaja de la aproximación bayesiana a la argumentación ha propiciado el desarrollo de un nuevo programa de investigación experimental que busca determinar en qué medida el desempeño de las personas en tareas de evaluación de argumentos informales se ajusta o no a las prescripciones bayesianas (Hahn y Oaksford 2007; 2012). En particular, los resultados de estos estudios empíricos sugieren que las personas evalúan argumentos de apelación a la ignorancia (*ad ignorantiam*), petición de principio (*petitio principii*), pendiente resbaladiza (Corner, Hahn y Oaksford 2011; Hahn y Oaksford 2007) y contra la persona (*ad hominem*) (Harris, Hsu y Madsen 2012) de manera consistente con los estándares propuestos por el modelo bayesiano. Finalmente, la evidencia recopilada durante las últimas dos décadas apunta al éxito de la aproximación bayesiana en la explicación de ciertas propiedades de otros procesos cognitivos complejos como percepción, aprendizaje, procesamiento lingüístico, categorización e inferencia causal (Chater y Oaksford 2008; Griffiths, Chater, Kemp, Perfors y Tenenbaum 2010).

En síntesis, Hahn y Oaksford (2012) señalan que el modelo bayesiano proporciona una suerte de “pegamento conceptual” que conecta y responde en el nivel teórico las preguntas en torno a la evaluación de la calidad de los argumentos que han surgido desde muy diversas tradiciones de investigación empírica en psicología. Sin embargo, desde la lógica informal y ciertos trabajos teóricos en el campo de la inteligencia artificial han surgido críticas con respecto a la idoneidad normativa y descriptiva del modelo bayesiano de actualización de creencias, las cuales también aplican para el caso de su análogo en evaluación de argumentos.

Críticas a la epistemología bayesiana como modelo normativo de la evaluación de argumentos. Pollock (2006) enumera algunas limitaciones de las aproximaciones bayesianas como modelo de la actualización de creencias. El primer problema surge del hallazgo empírico de que las personas no realizan asignaciones de probabilidades coherentes con los axiomas del cálculo de probabilidades, sino que exhiben desviaciones sistemáticas con respecto a ellos. Es así como

en la investigación sobre juicios probabilísticos y toma de decisiones se ha encontrado que las personas asignan una probabilidad a la disyunción de eventos exhaustivos y mutuamente excluyentes menor (subaditiva) o mayor (súperaditiva) que la suma total de las probabilidades de cada evento por separado (Tversky y Koehler 1994) o consideran más probable la ocurrencia conjunta de ciertos eventos que su ocurrencia aislada, incurriendo en la denominada “falacia de la conjunción” (Tversky y Kahneman 1983). Aun suponiendo que la probabilidad subjetiva se identifica con los grados de creencia que las personas deberían tener y no con los que de hecho tienen, el modelo bayesiano no prescribe con exactitud qué tipo de ajuste se debe realizar para convertir las asignaciones inconsistentes en la asignación “racional”. Para ilustrar este punto, Pollock plantea el caso de una persona que realiza la siguiente asignación de probabilidades: $P(A) = 0.5$, $P(B) = 0.7$, $P(A \vee B) = 0.8$, $P(A \wedge B) = 0.3$. Este conjunto es inconsistente con los axiomas de Kolmogórov, dado que $P(A \vee B) \neq P(A) + P(B) \mid P(A \wedge B)$. Para restaurar la coherencia con dichos axiomas (y con ello la racionalidad de la asignación), hay un número potencialmente infinito de modificaciones posibles que pueden realizarse sobre el conjunto, tales como:

- 1) $P(A) = 0.4 \quad P(B) = 0.7 \quad P(A \vee B) = 0.8 \quad P(A \wedge B) = 0.3$
- 2) $P(A) = 0.5 \quad P(B) = 0.6 \quad P(A \vee B) = 0.8 \quad P(A \wedge B) = 0.3$
- 3) $P(A) = 0.5 \quad P(B) = 0.7 \quad P(A \vee B) = 0.9 \quad P(A \wedge B) = 0.3$
- 4) $P(A) = 0.5 \quad P(B) = 0.7 \quad P(A \vee B) = 0.8 \quad P(A \wedge B) = 0.4$

El “problema de la unicidad” radica en que, una vez se satisface el requerimiento de consistencia, no es posible definir a partir del cálculo de probabilidades un único conjunto “racional” de grados de creencia que reemplace la asignación inconsistente. En consecuencia, la posibilidad de establecer grados de creencia “racionales” depende crucialmente de que las asignaciones de probabilidades previas sean coherentes con los axiomas de Kolmogórov, lo cual puede no ocurrir en las inferencias cotidianas, según ha mostrado la evidencia empírica. Si una persona adjudica grados de creencia previos en las premisas y conclusión de un argumento que sean inconsistentes con el cálculo de probabilidades, el modelo de evaluación bayesiano no puede predecir con exactitud su grado de creencia posterior en la conclusión.

En segundo lugar, debido a que la probabilidad de la conjunción es siempre inferior a la probabilidad de sus elementos, esto es, $P(A \wedge B) < P(A)$, ($P(B)$ cuando $P(A | B) < 1$ y $P(B | A) < 1$, a medida que se introducen más premisas en un argumento la probabilidad de la conclusión decrece, de manera que éste puede debilitarse tanto como se quiera a partir de la sola adición de premisas (Pollock 2008). Por ejemplo, supóngase que un investigador realiza una encuesta por internet a 1000 personas y determina que 790 de ellas manifiestan no estar de acuerdo con el patrocinio del estado a las corridas de toros. Sin embargo, el investigador es plenamente consciente de que algunas personas pudieron haber indicado por error en la encuesta la opción que no correspondía con su opinión, de manera que asigna una probabilidad de 0.999 a la respuesta de cada encuestado. De acuerdo con el cálculo de probabilidades, el grado de creencia que el investigador debería asignar a la conclusión “El 79 por ciento de las personas encuestadas no está de acuerdo con que las corridas de toros sean patrocinadas por el estado” es de $0.999^{1000} = 0.37$. En el ámbito de la evaluación de argumentos, resulta paradójico que un número elevado de premisas altamente probables conduzca a una disminución en la probabilidad de la conclusión.

Las dos objeciones mencionadas cuestionan la pertinencia de representar el grado de creencia de las personas en términos de probabilidades subjetivas consistentes con las restricciones del cálculo de probabilidades (en particular, el requerimiento de asignación coherente de probabilidades y la definición de probabilidad conjunta), en la medida en que algunos de estos supuestos impiden capturar formas intuitivamente legítimas de inferencia. Rescher (1976) ilustra un resultado contraintuitivo similar que se sigue de la aplicación del cálculo de probabilidades a un esquema inferencial más cotidiano: considérese una cuadrícula con 9 posiciones posibles, tal como la que se presenta en la figura 1. Supongamos que se sabe que existe un objeto z que se encuentra dentro de la cuadrícula, aunque su posición real no se conoce con certeza, y que dos fuentes, X e Y , se consideran moderadamente confiables en la medida en que $P(A | X \text{ reporta } A) = P(B | Y \text{ reporta } B) = 0.7$, para todo A y B que sea declarado por alguna de las dos fuentes. Si X declara que “el objeto z se encuentra en la columna 2”, mientras que Y declara que “el objeto z se encuentra en la fila B”, la distribución de las probabilidades de que el objeto z se encuentre en cada una de las celdas de acuerdo con ambos testimonios puede ser representada a través de la figura 1.

A partir de esta distribución, se puede inferir que la probabilidad de que el objeto z se ubique en la posición $B2$ equivale a $P(B2 |$

reportes de X e Y) = 0.49, mientras que la probabilidad de que se ubique en cualquier otra posición de la cuadrícula es igual a $P(\neg B2 | \text{reportes de } X \text{ e } Y) = 1 - P(B2 | \text{reportes de } X \text{ e } Y) = 0.51$. De manera paradójica, la probabilidad de que los testimonios de dos fuentes moderadamente confiables sean verdaderos al mismo tiempo es menor que la probabilidad de que no lo sean.

	1	2	3	
	0.15	0.7	0.15	
A	0.1	0.23	0.105	0.23
B	0.7	0.105	0.49	0.105
C	0.1	0.23	0.105	0.23

Figura 1. *Distribución de probabilidades de la presencia de un objeto en función del testimonio de dos fuentes moderadamente confiables (adaptado de Rescher 1976, p. 30).*

Tomando en consideración los resultados paradójicos mencionados, Rescher (1976) y posteriormente Walton (2008) defienden la existencia de un modelo inferencial alternativo (*i.e.*, el razonamiento plausible) que da cuenta de las inferencias propias de escenarios que no pueden ser modelados ni deductiva ni probabilísticamente. En consecuencia, Walton *et al.* (2013) postulan que “el razonamiento plausible admite grados de prueba, pero de un tipo diferente de los valores de probabilidad estándar y reglas bayesianas usados en la probabilidad pascaliana” (p. 12). Mientras que los proponentes del razonamiento plausible enfatizan en que sus propiedades le permiten superar las limitaciones del razonamiento probabilístico (Rescher 1976; Walton 2008), los defensores del modelo bayesiano sostienen que la postulación de un “tercer” tipo de estándar es innecesaria y que la aplicación de las reglas del razonamiento plausible a la evaluación de argumentos puede conducir a resultados contraintuitivos (Hahn y Oaksford 2006; Hahn *et al.* 2013). En el apartado siguiente, se presentarán los principios básicos del razonamiento plausible como modelo formal de evaluación de argumentos, a partir de las aproximaciones teóricas de Rescher y Walton.

¿Qué es el razonamiento plausible? En su trabajo seminal sobre el tema, Rescher (1976) señala que el desarrollo de la teoría del razonamiento plausible obedece a dos propósitos principales: 1) sistematizar las inferencias que se realizan a partir de enunciados cuyo único respaldo es la “confiabilidad”, “credibilidad” o “solidez” de las fuentes que las declaran y 2) proporcionar herramientas formales que permitan realizar inferencias en situaciones de inconsistencia informacional (*i.e.*, aquellas en las que reportes de distintas fuentes se contradicen entre sí), dado que la lógica deductiva y el cálculo de probabilidades proscriben la inferencia a partir de contradicciones. Así, “el punto de la teoría de plausibilidad es proporcionar un medio para operar con los grados relativos de aceptabilidad o presunción en favor de datos diversos” (Rescher 1976, p. 10).

Desde esta aproximación, el concepto de fuente no sólo hace referencia a agentes (*e.g.*, expertos, testigos, autoridades, instituciones) u objetos concretos particulares (*e.g.*, registros oficiales, libros, bases de datos en general), sino también a ciertos productos de procesos inferenciales (*e.g.*, teorías, conjeturas, suposiciones, principios, reglas, tradiciones, etc.). En el razonamiento plausible, la confiabilidad de la fuente y la plausibilidad de sus declaraciones están directamente relacionadas (a menos que se especifique lo contrario): “Mientras más grande sea la confiabilidad de la fuente, mayor será la plausibilidad (en este sentido) con la cual ésta puede asegurar las tesis que declara. En general —aunque no invariablemente— estos dos hechos serán paralelos entre sí” (Rescher 1976, p. 11).

La plausibilidad de un conjunto de proposiciones declaradas por una fuente puede ser expresada en una escala en la que 1 equivale a la plausibilidad absoluta y la serie $S_n^{\frac{n}{n}}, n - \frac{1}{n}, n - \frac{2}{n}, n - \frac{3}{n} \dots \frac{1}{n}$, representa los distintos grados de plausibilidad que pueden ser atribuidos a dicho conjunto, de manera que éste no podrá tener una plausibilidad nula o negativa. Teniendo en cuenta que la plausibilidad se define en una escala *ordinal*, Rescher aclara que el propósito de esta formalización no es definir un cálculo cuantitativo análogo al cálculo de probabilidades, sino más bien proporcionar un cálculo “cuantitativo” más simple, que permita establecer la igualdad o desigualdad entre los valores de plausibilidad de diferentes conjuntos de proposiciones. En consecuencia, ciertas restricciones del cálculo de probabilidades no aplican para el razonamiento plausible, tal y como se hace evidente en los axiomas propuestos por Rescher para este último:

- a. Condición de metrización (*Metrization condition*): la plausibilidad de una proposición sólo puede tomar valores entre 0 y 1.

Simbólicamente, para toda proposición A , existe un valor k tal que $0 < k < 1$, tal que $\text{Plaus}(A) = k$.

- b. Condición de maximización de la verdad lógica: las tautologías tienen el máximo valor de plausibilidad posible. Si A es una tautología, entonces $\text{Plaus}(A) = 1$.
- c. Condición de compatibilidad: todas las proposiciones a las que se les asigne el máximo valor de plausibilidad posible deben ser lógica y fácticamente consistentes. Para todo A y B tal que $\text{Plaus}(A) = \text{Plaus}(B) = 1$, se cumple que $\neg((A \wedge \neg B) \vee (\neg A \wedge B))$.
- d. Condición de consecuencia: una proposición derivada de un conjunto de proposiciones consistentes entre sí debe ser al menos tan plausible como la premisa menos plausible del conjunto. Si C es un conjunto de proposiciones consistentes, $C = A_1 \dots A_n$ y $A_1 \dots A_n \vdash B$, entonces $\text{Plaus}(\min(A_1 \dots A_n)) \leq \text{Plaus}(B)$.
- e. Estipulación de inconsistencia: la plausibilidad de una afirmación y de su negación pueden ser altamente plausibles de manera simultánea (e.g., $\text{Plaus}(A) = \text{Plaus}(\neg A) = 0.9$).

De los axiomas precedentes se deducen los siguientes teoremas:

1. Si $A \vdash B$ entonces $\text{Plaus}(A) \leq \text{Plaus}(B)$.
2. Si $A \vdash B$ y $B \vdash A$, entonces $\text{Plaus}(A) = \text{Plaus}(B)$.
3. Regla MIN o de Teofrastro: para todo A y B , se cumple que $\text{Plaus}(A \wedge B) = \min(\text{Plaus}(A), \text{Plaus}(B))$.
4. Regla MAX: para todo A y B , $\text{Plaus}(A \vee B) \geq \max(\text{Plaus}(A), \text{Plaus}(B))$.

A partir de lo anterior, es posible indicar al menos tres diferencias claras entre las prescripciones bayesianas y las del razonamiento plausible:

- 1) La estipulación de inconsistencia contradice directamente la ley de negación del cálculo de probabilidades, según la cual $P(A) + P(\neg A) = 1$.
- 2) La regla MIN es inconsistente con la regla de la conjunción del cálculo de probabilidades, de acuerdo con la cual $P(A \wedge B) \leq P(A), P(B)$.

- 3) La regla MAX es incompatible con una fórmula ampliamente utilizada para calcular la probabilidad de un efecto dado un conjunto de causas independientes en una red bayesiana, denominada NOISY- OR (Pearl 1988), donde p_i representa la probabilidad condicional de que la conclusión C sea verdadera dado que la premisa x_i es verdadera y las demás premisas son falsas:

$$P(C | x_1 \dots x_i) = 1 - \prod_{i:x_i \in X_p} (1 - p_i)$$

Las marcadas diferencias entre los cálculos probabilístico y de plausibilidad podrían en principio hacer pensar que ambos tipos de inferencia tienen poco en común, dado que de ellos se derivan predicciones divergentes cuando se usan como modelos normativos de argumentación.⁶ No obstante, el análisis realizado en la primera sección ha mostrado que los conceptos de probabilidad subjetiva y están estrechamente relacionados⁷ y que las características que ambos comparten son al menos tan importantes como aquellas que los distinguen.

La tabla 1 resume el análisis precedente, muestra las características compartidas y no compartidas por las distintas formas de inferencia examinadas.

Aunque distintos teóricos enfatizan la incompatibilidad entre la teoría de la probabilidad y la inferencia plausible (Pollock 2006; Rescher 1976; Walton 2008), estos autores parecen tomar distancia con respecto a una visión frecuentista fuerte de la idea de probabilidad, en la que el criterio evaluativo es siempre externo y existe un compromiso con una noción de verdad basada en la correspondencia (*i.e.*, entre la estimación y las regularidades estadísticas que ocurren en el mundo). En contraste, las relaciones conceptuales entre la idea de plausibilidad y la interpretación subjetiva de la probabilidad no son examinadas en detalle. Adicionalmente, la ausencia de un cálculo lógico de la plausibilidad comparable al cálculo de probabilidades restringe la comparación entre ambos esquemas inferenciales a señalar las discrepancias en ciertas reglas.

⁶ Este tema lo trataremos más adelante.

⁷ Sobre la relación histórica de estos conceptos, véase el capítulo 1 de Walton 1998.

TABLA 1. *Comparación entre tipos de inferencia*

Propiedad	Tipo de inferencia			
	Deductiva	Probabilística (frecuentística)	Probabilística (Subjetivista)	Plausibilidad
Relatividad a la audiencia	×	×	✓	✓
No-monotonía	×	✓	✓	✓
Cuantificación de la fuerza de un argumento	×	✓	✓	✓
Independencia del concepto de verdad	×	×	✓	✓
Posibilidad de inferencias a partir de contradicciones	×	×	×	✓

Más allá de los “parecidos de familia” que emparentan los conceptos de plausibilidad y probabilidad subjetiva, el razonamiento plausible definido que define Rescher y retoma Walton es un modelo alternativo de ciertos tipos de inferencia que sólo pueden ser interpretadas desde la aproximación bayesiana como desviaciones normativas. Si esto es cierto desde el punto de vista descriptivo, deben existir tareas de evaluación de argumentos en las que sea posible predecir correctamente los desempeños de las personas a partir del modelo de razonamiento plausible, pero no del modelo bayesiano.

4. Probabilidad, plausibilidad y falacias

Dejando de lado por un momento las “declaraciones de principios” en el nivel teórico, una manera alternativa de aproximarse al problema de la distinción entre plausibilidad y probabilidad subjetiva consiste en comparar las implicaciones que cada modelo puede tener con respecto a un mismo fenómeno argumentativo. Recientemente, la aproximación bayesiana a la argumentación se ha ocupado de 1) establecer los criterios que permiten distinguir entre argumentos falaces y razonables desde un punto de vista probabilístico y 2) determinar si tales criterios coinciden o no con los que emplean de manera implícita las personas a la hora de juzgar un argumento como razonable o falaz en escenarios cotidianos. En particular, estas aproximaciones han aplicado esta estrategia al análisis de las falacias de apelación a la ignorancia (*ad ignorantiam*), petición de principio (*petitio principii*) (Hahn y Oaksford 2007), pendiente resbaladiza (Corner *et al.* 2011) y, más recientemente, argumentos contra la persona (*ad hominem*) (Harris *et al.* 2012). En este apartado se presentarán brevemente tres falacias: apelación a la autoridad (*ad verecundiam*), a la popularidad (*ad populum*) y petición de principio, examinando las diferencias y semejanzas entre los criterios normativos que pueden derivarse desde la aproximación bayesiana y desde el razonamiento plausible. Adicionalmente, se presentarán algunos hallazgos empíricos relacionados, provenientes de estudios recientes en ciencia cognitiva sobre argumentación.

Apelación a la autoridad. En los escenarios argumentativos cotidianos, las personas suelen utilizar los dictámenes de expertos para dar apoyo a sus conclusiones, desde el diagnóstico médico o el peritaje forense, hasta la verificación de la autenticidad de una obra de arte o la planeación de políticas públicas. Debido al actual grado de especialización en la producción del conocimiento, un agente que pretenda recolectar y analizar por su cuenta toda la evidencia pertinente

para sustentar un determinado argumento enfrentaría restricciones insalvables de tiempo y recursos. En este sentido, la apelación a la autoridad de expertos es una suerte de mecanismo de “división del trabajo cognitivo” que permite el acopio de evidencia proveniente de distintas fuentes especializadas en su obtención (Woods y Walton 1982). Sin embargo, como resulta obvio, no todas estas apelaciones son igualmente razonables. Por ejemplo, es cuestionable citar a Einstein como autoridad sobre temas políticos basándose en sus logros en el campo de la física teórica. Copi y Cohen (2005) definen la falacia *ad verecundiam* como “una apelación a personas que no tienen credenciales legítimas de autoridad en materia de discusión” (p. 130). Por otro lado, también sería impensable que el dictamen de un prestigioso laboratorio sobre la presencia de ADN de un sospechoso en la escena de un crimen sea tomado en cuenta si las muestras no fueron recolectadas siguiendo un protocolo riguroso. Para que un argumento de autoridad sea sólido, la fuente citada debe ser capaz de proporcionar evidencia robusta y relevante que lo sustente cuando ésta sea requerida (Woods y Walton 1982).

¿Qué rasgos distinguen una apelación racional al testimonio de un experto de una falaz? Walton (2006) postula seis preguntas críticas que pueden utilizarse para evaluar diferentes aspectos de los cuales depende la fuerza de un argumento que apela a la opinión de expertos: 1) *Experticia* (¿Qué tan experta es la fuente en el tema de la discusión?). 2) *Relevancia* (¿Qué tan relacionado está el contenido del argumento con el campo de experticia de la fuente?). 3) *Opinión* (¿Qué tanto apoyo confiere la fuente a la conclusión?). 4) *Fiabilidad* (¿Qué tan confiable es la fuente?). 5) *Consistencia* (¿Qué tan consistente es la opinión de este experto con la de otros expertos?). 6) *Respaldo en la evidencia* (¿Qué tanto se basa la opinión del experto en evidencia objetiva?). Hahn *et al.* (2013) señalan que la respuesta a cada una de estas seis preguntas puede precisarse dentro una red bayesiana. Desde esta perspectiva, las respuestas a las preguntas 1) y 2) serían tratadas como nodos distintos que dan soporte a la opinión de la fuente experta, la cual a su vez podría ser apoyada por evidencia adicional proporcionada por dicha fuente (pregunta 6). Nodos adicionales en la red permitirían capturar el grado de consistencia existente entre el dictamen de la fuente y el de otros expertos (pregunta 5).

En contraste, Woods y Walton (1982) señalan que la apelación a testimonios expertos sólo puede ser coherente con la perspectiva probabilística cuando todas las fuentes consultadas coinciden en sus evaluaciones de la probabilidad de que la conclusión que se pretende establecer sea verdadera. En caso de que una fuente asigne un valor

de probabilidad totalmente distinto al que otra fuente igualmente confiable asigna a la misma proposición, los axiomas de la teoría de la probabilidad restringen la posibilidad de realizar inferencias. Esto se debe a que la probabilidad condicional de una afirmación dada una contradicción (en este caso, dos testimonios que se desmienten entre sí) equivale a una división por 0 y, por tanto, es indefinida.

En la misma línea de Rescher (1976), Woods y Walton defienden que el razonamiento plausible provee una alternativa para evaluar argumentos construidos a partir de conjuntos de premisas inconsistentes. Desde este punto de vista, la fuerza de la apelación a un testimonio de expertos debe ser evaluada en función de dos factores: la confiabilidad de la fuente y la plausibilidad de lo que ésta afirma. Ambos términos conforman un argumento ligado, en el que el debilitamiento de cualquiera de los dos restaría fuerza a todo el argumento. Específicamente, en este tipo de argumentos, la conclusión no puede ser más plausible que la menos plausible de sus premisas (regla MIN o de Teofrastro) (Walton 2008). Así, si una fuente X con una confiabilidad atribuida de 0.8 declara que p , una proposición cuyo valor de plausibilidad previa es de 0.2, la plausibilidad de todo el argumento no puede ser superior a este último valor. Sin embargo, Hahn, Harris y Corner (2009; Exp. 1) presentan evidencia de que las personas no ponderan la confiabilidad de la fuente y la cantidad de evidencia disponible de acuerdo con esta regla, sino de manera multiplicativa, en consonancia con las prescripciones del modelo bayesiano. Aunque este hallazgo constituye un desafío para las prescripciones del modelo de argumentación plausible de Walton (2008), Hahn y sus colaboradores no incluyen en este estudio una condición experimental en la que los participantes se vean enfrentados a un disenso entre fuentes de igual o distinta confiabilidad, lo cual constituye el escenario particular en el que Woods y Walton señalan que las reglas del razonamiento plausible resultan más apropiadas que las probabilísticas.

Apelación a la popularidad. Un contexto argumentativo con características similares a las del anterior es la apelación a la opinión popular, en el cual el hecho de que la mayoría de personas acepta cierta proposición p como verdadera se invoca como una razón en favor de p (Walton 2006). Al igual que en el caso de la apelación a testimonios expertos, existen usos falaces de apelación a la opinión popular (*i.e.*, *falacia ad populum*). Por ejemplo, concluir que la marihuana es un producto recomendable para todos basándose en que millones de personas en el mundo la consumen a diario. Ahora bien, también existen casos en los que el hecho de que la mayoría esté de acuerdo en la verdad de una proposición puede ser un criterio

razonable para aumentar su credibilidad. Por ejemplo, el hecho de que varios testigos presentes en la escena del crimen identifiquen a cierto sospechoso como el culpable constituye una evidencia más robusta de su culpabilidad que el caso en el que sólo un testigo logra identificarlo.

Harris y Hahn (2009) muestran que las personas son sensibles a la coherencia entre testimonios provenientes de fuentes independientes y que son capaces de combinar esta información con la confiabilidad atribuida a cada testigo, de acuerdo con las prescripciones del modelo bayesiano. En particular, las personas juzgan más fuerte un argumento basado en testimonios múltiples cuando más testigos coinciden, aun cuando su confiabilidad individual sea la misma. Desde el modelo de plausibilidad, un argumento basado en múltiples testimonios independientes constituye un argumento convergente, en el cual la negación de uno de los testimonios no disminuye significativamente el apoyo que los demás brindan a la conclusión. De acuerdo con Walton (1992 y 2008), en este tipo de argumentos la plausibilidad de la conclusión debe ser al menos tan alta como la más plausible de las premisas que la sustentan.⁸ Así, si se tienen dos testimonios que declaran la ocurrencia de un mismo hecho, uno con una plausibilidad atribuida de 0.6 y otro con una de 0.8, el límite inferior de la plausibilidad de la conclusión sería este último valor. Desde este punto de vista, recopilar testimonios adicionales con una plausibilidad inferior a 0.8 no incrementaría en principio la plausibilidad del argumento. No obstante, Walton y Reed (2008) reconocen que la adición de testimonios en favor de una misma conclusión sí aumenta la plausibilidad de dicha conclusión, en la medida en la que elimina una posible refutación del argumento que podría haber provenido de tal fuente. Consecuentemente, la introducción de un testimonio que refute la conclusión hace más débil el argumento. Sin embargo, ni Walton ni Rescher proporcionan una regla de evaluación numérica que permita cuantificar la magnitud de este incremento en función de la confiabilidad de la fuente o del número de testimonios coherentes. En este sentido, no resulta claro qué tan distintas pueden ser las prescripciones normativas del modelo de plausibilidad con respecto al tratamiento bayesiano en este escenario argumentativo.

Petición de principio. Un tercer tipo de falacia particularmente interesante para la contrastación teórica y empírica de ambos modelos es la petición de principio. Ésta se define como un argumento en el cual la conclusión que se intenta establecer se asume implícita

⁸ Véase anteriormente la regla MAX.

o explícitamente como premisa o en el que la verdad de una de las premisas depende de que la conclusión sea verdadera (Copi y Cohen 2005). Así, en el argumento “Dios existe porque la Biblia lo dice y la Biblia es la palabra de Dios”, la verdad de la premisa “La Biblia es la palabra de Dios” depende crucialmente de la verdad de la conclusión “Dios existe”. Este argumento tiene una estructura similar al siguiente: “El bosón de Higgs existe porque podemos ver la traza x en el acelerador de partículas y la traza x es la marca que deja el bosón de Higgs.” No obstante, la comunidad científica estaría dispuesta a aceptar el segundo argumento pero no el primero.⁹ ¿Qué hace a un argumento más razonable que a otro? Hahn y Oaksford (2007) abordan esta pregunta diferenciando tres componentes en este tipo de inferencias: la evidencia encontrada (*e.g.*, la Biblia, la traza X), la interpretación de la evidencia (*e.g.*, “La Biblia es la palabra de Dios”, “la traza x es la marca que deja el bosón de Higgs”) y la conclusión del argumento (*e.g.*, “Dios existe”, “El bosón de Higgs existe”). En el caso de una petición de principio, la interpretación de la evidencia sólo puede ser verdadera cuando la conclusión también lo es. De tal modo que si la conclusión tiene una probabilidad previa baja, su interpretación también la tendrá. En general, Hahn y Oaksford establecen que en este tipo de argumentos la probabilidad de la interpretación de la evidencia no puede ser más alta que la probabilidad de la conclusión. Sin embargo, cuando la evidencia es encontrada (*e.g.*, la traza x es visible) y su interpretación dentro del argumento es probablemente correcta (*e.g.*, la probabilidad de que la traza x se deba a un fenómeno distinto al bosón de Higgs es baja), la petición de principio no es falaz.

En contraste, Walton (1992) propone que los argumentos deductivamente válidos (como es el caso de una petición de principio) son de tipo ligado. Como se vio anteriormente, desde la perspectiva del razonamiento plausible este tipo de argumentos deben evaluarse siguiendo la regla según la cual la conclusión puede ser a lo sumo tan plausible como la más plausible de las premisas. Sin embargo, si la conclusión es en sí misma parte de las premisas, es imposible que la plausibilidad del argumento sea superior a la plausibilidad previamente atribuida a la conclusión. En otras palabras, un argumento que

⁹ Por supuesto, se puede argumentar que la similitud entre los argumentos es puramente superficial, dado que el caso del bosón puede caracterizarse como una inferencia abductiva o “inferencia a la mejor explicación” a partir de la traza, mientras que el de la existencia de Dios no puede ser “explicada” acudiendo a la Biblia. La similitud superficial, sin embargo, es suficiente para el punto que se expone a continuación.

involucre una petición de principio en ningún caso puede incrementar la plausibilidad de la conclusión. Hahn y Oaksford (2006) califican este tratamiento como contraintuitivo, dado que muchas “inferencias a la mejor explicación” (Hanson 1971, y Lipton 2004) dentro de las ciencias (como el mencionado ejemplo del bosón de Higgs) poseen una estructura similar a la de una petición de principio. Adicionalmente, los hallazgos empíricos de Hahn y Oaksford (2007) sugieren que las personas toman en cuenta la cantidad de interpretaciones alternativas posibles de la evidencia a la hora de juzgar la fuerza de un argumento con la estructura de una petición de principio. En particular, mientras menor es el número de interpretaciones posibles de la evidencia que no impliquen la verdad de la conclusión, el argumento será mejor evaluado.

5. Conclusión: entre lo descriptivo y lo normativo

A diferencia de lo que ocurre con los modelos bayesianos de evaluación de argumentos, no existe hasta el momento algún resultado empírico que sugiera que las personas evalúan argumentos de acuerdo con las reglas prescritas por el razonamiento plausible. Desde una perspectiva más amplia, el presente análisis sugiere que los contextos argumentativos con las condiciones adecuadas para la contrastación entre modelos probabilísticos y de plausibilidad pueden ser menos comunes de lo que podría pensarse en un principio. En los tres escenarios argumentativos examinados, se encontró que la aplicación de las normas del razonamiento plausible efectivamente puede conducir o bien a resultados contrarios a la intuición (*e.g.*, petición de principio), a prescripciones poco precisas (*e.g.*, apelaciones a la popularidad) o que no coinciden con la manera en la cual las personas evalúan argumentos (*e.g.*, apelaciones a la autoridad).

Recientemente, los autores Gaviria y Jiménez-Leal (en prensa) han abordado experimentalmente la comparación del ajuste descriptivo del modelo bayesiano de evaluación de argumentos y del razonamiento plausible. Los resultados encontrados sugieren que el modelo bayesiano de argumentación describe mejor el desempeño de las personas en tareas de evaluación de argumentos cotidianos y de apelaciones testimoniales que el modelo de razonamiento plausible. En el caso de los argumentos cotidianos, se encontró que el patrón de respuestas parece ser opuesto al que prescriben las reglas MAX y MIN del razonamiento plausible: las personas tienden a evaluar los argumentos convergentes como menos fuertes que la premisa más fuerte y los argumentos ligados como más fuertes que la premisa

menos fuerte. En el caso de las apelaciones testimoniales, la fuerza percibida de los argumentos fue inferior a la del testimonio más confiable, tanto en estructuras convergentes como ligadas. Por otro lado, los juicios evaluativos de las personas que identifican un determinado argumento cotidiano como convergente exhiben un mayor ajuste a las predicciones del modelo bayesiano que a las del razonamiento plausible, mientras que sólo el modelo bayesiano da cuenta de los efectos combinados de coherencia y estructura observados en las evaluaciones de apelaciones testimoniales.

El hecho de que el modelo bayesiano haya demostrado ser descriptivamente más adecuado que el razonamiento plausible no significa que el primero sea *per se* un “mejor” marco normativo que el segundo. Lo que este hecho sugiere es que el modelo bayesiano permite representar de manera más fiel las propiedades de las tareas que realizan las personas al evaluar cierto tipo de argumentos cotidianos y de apelaciones testimoniales. Corner y Hahn (2013) proporcionan una analogía ilustrativa sobre este punto, proveniente del ámbito legal: usar una tarjeta de crédito de otra persona sin su permiso puede ser tipificado como robo o fraude; la decisión de aplicar a este caso o bien las normas del código penal concernientes al robo o bien las referidas al fraude no afecta en modo alguno el estatus legal de dichas normas. En el caso de la argumentación, la comparación entre la conducta de los participantes y un determinado modelo normativo permite establecer hasta qué punto la gran variedad de tareas de evaluación de argumentos que las personas realizan cotidianamente pueden ser entendidas o no en términos de dicho modelo (Corner y Hahn 2013). Naturalmente, variaciones tanto en las metas como en las restricciones ambientales y de procesamiento de dichas tareas pueden hacer que el modelo normativo correspondiente también cambie.

Una alternativa relevante al respecto del modelo normativo la constituye el llamado razonamiento conductivo (Pinto 2011). El razonamiento conductivo es un tipo de inferencia revocable que captura algunas de las características del razonamiento plausible. Por ejemplo, en el contexto de este tipo de razonamiento, la fuerza de una premisa es una función del riesgo de asumirla y de su peso relativo en un argumento (Pinto 2011). Nociones como éstas son muy difíciles de recoger en un marco normativo como el probabilista, que se caracteriza por la descripción unidimensional de la fuerza del argumento. Ideas como la del razonamiento conductivo todavía están por ser exploradas en ciencia cognitiva.

En dos pasajes particularmente oscuros, Walton señala a propósito de la relación entre razonamiento plausible y probabilidad subjetiva:

Muchos sostienen que la estructura lógica del razonamiento plausible puede reducirse a alguna forma de razonamiento inductivo (basado en probabilidad). Y ésta no es una afirmación que me encargaría de rechazar categóricamente. Está por verse si la reducción puede llevarse a cabo. (Walton 2002, p. 111)

El razonamiento plausible de este tipo es muy común, pero bastante diferente del razonamiento sobre la base de probabilidades del tipo modelado en el cálculo de probabilidades. Por supuesto, podría representar algún tipo de probabilidad subjetiva. (Walton 2008, p. 69)

Reexaminar hasta donde se extienden las afinidades conceptuales entre las dos formas de inferencia y si es necesario o no proponer la existencia de un “tercer tipo” de modelo de evaluación de argumentos es una cuestión abierta que pide respuestas tanto de la lógica informal como de la ciencia cognitiva de la argumentación.

BIBLIOGRAFÍA

- Aristóteles, 1994, *Metafísica*, trad. T. Calvo Martínez, Gredos, Madrid.
- _____, 1988, *Tratados de Lógica (Órganon)*, vol. II, trad. M. Candel Sanmartín, Gredos, Madrid.
- _____, 1982, *Tratados de Lógica (Órganon)*, vol. I, trad. M. Candel Sanmartín, Gredos, Madrid.
- Blair, J.A., 2011, *Groundwork in the Theory of Argumentation: Selected Papers of J. Anthony Blair*, Springer, Dordrecht.
- Chater, N. y M. Oaksford, 2008, *The Probabilistic Mind: Prospects for Bayesian Cognitive Science*, Oxford University Press, Oxford, Reino Unido.
- Copi, I. y C. Cohen, 2005, *Introducción a la lógica*, Limusa, México.
- Corner, A. y U. Hahn, 2013, “Normative Theories of Argumentation: Are Some Norms Better than Others?”, *Synthese*, vol. 190, no. 16, pp. 3579–3610.
- Corner, A., U. Hahn y M. Oaksford, 2011, “The Psychological Mechanism of the Slippery Slope Argument”, *Journal of Memory y Language*, vol. 64, no. 2, pp. 133–152.
- Gaviria, C. y W. Jiménez-Leal, en prensa, “Coherence and Argument Structure: An Empirical Comparison between Plausible Reasoning and the Bayesian Approach to Argumentation”, en P. Bello, M. Guarini, M. McShane, y B. Scassellati (comps.), *Proceedings of the 36th Annual*

*Queremos agradecer de manera especial los comentarios de un árbitro anónimo, cuya lectura cuidadosa nos permitió detectar y corregir algunos errores. Para la realización de este proyecto William Jiménez-Leal recibió financiamiento del fondo FAPA (Fondo para profesores Asistentes) de la Universidad de los Andes.

- Conference of the Cognitive Science Society*, Cognitive Science Society, Austin, Texas.
- Griffiths, T.L., N. Chater, C. Kemp, A. Perfors y J. Tenenbaum, 2010, “Probabilistic Models of Cognition: Exploring Representations and Inductive Biases”, *Trends in Cognitive Sciences*, vol. 14, no. 8, pp. 357–364.
- Hahn, U., A.J.L. Harris, y A. Corner, 2009, “Argument Content and Argument Source: An Exploration”, *Informal Logic*, vol. 29, no. 4, pp. 337–367.
- Hahn, U. y M. Oaksford, 2012, “Rational Argument”, en K.J. Holyoak y R.G. Morrison (comps.), *The Oxford Handbook of Thinking and Reasoning*, Oxford University Press, Nueva York.
- _____, 2007, “The Rationality of Informal Argumentation: A Bayesian Approach to Reasoning Fallacies”, *Psychological Review*, vol. 114, no. 3, pp. 704–732.
- _____, 2006, “A Normative Theory of Argument Strength”, *Informal Logic*, vol. 26, no. 1, pp. 1–22.
- Hahn, U., M. Oaksford, y A.J.L. Harris, 2013, “Testimony and Argument: A Bayesian Perspective”, en F. Zenker (comp.), *Bayesian Argumentation*, Springer, Dordrecht.
- Hanson, N.R., 1971, *Observación y explicación: guía de la filosofía de la ciencia; Patrones de descubrimiento: investigación de las bases conceptuales de la ciencia*, Alianza, Madrid.
- Harris, A.J.L. y U. Hahn, 2009, “Bayesian Rationality in Evaluating Multiple Testimonies: Incorporating the Role of Coherence”, *Journal of Experimental Psychology: Learning, Memory, and Cognition*, vol. 35, pp. 1366–1372.
- Harris, A.J.L., A.S. Hsu, y J.K. Madsen, 2012, “Because Hitler Did It! Quantitative Tests of Bayesian Argumentation Using ‘ad hominem’”, *Thinking and Reasoning*, vol. 18, no. 3, pp. 311–343.
- Lipton, P., 2004, *Inference to the Best Explanation*, 2a. ed., Routledge, Londres.
- Pearl, J., 1988, *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*, Morgan Kauffman Publishers, San Mateo, California.
- Pinto, R., 2011, “Weighing Evidence in the Context of Conductive Reasoning”, en J. Blair y R. Johnson (comps.), *Conductive Argument. An Overlooked Type of Defeasible Reasoning*, College Publications, Londres.
- Pollock, J.L., 2008, “Defeasible Reasoning”, en L.J. Rips, y J. Adler (comp.), *Reasoning: Studies in Human Inference and Its Foundations*, Cambridge University Press, Cambridge, Mass.
- _____, 2006, *Thinking about Acting: Epistemological Foundations for Rational Decision Making*, Oxford University Press, Nueva York.

- Pollock, J.L., 1995, *Cognitive Carpentry: A Blueprint for How to Build a Person*, The MIT Press, Cambridge, Mass.
- , 1992, "How to Reason Defeasibly", *Artificial Intelligence*, vol. 57, pp. 1–42.
- Real Academia Española, 2001, Diccionario de la lengua española (22 ed.) en <http://www.rae.es/rae.html> [fecha de consulta: mayo 2014].
- Rescher, N., 1976, *Plausible Reasoning: An Introduction to the Theory and Practice of Plausibilistic Inference*, Van Gorcum, Assen.
- Tversky, A. y D. Kahneman, 1983, "Extensional Versus Intuitive Reasoning: The Conjunction Fallacy in Probability Judgment", *Psychological Review*, vol. 90, no. 4, pp. 293–315.
- Tversky, A. y D. Koehler, 1994, "Support Theory: A Non-Extensional Representation of Subjective Probability", *Psychological Review*, vol. 101, no. 4, pp. 547–567.
- Vega-Reñón, L., 1998, "Aristotle's Endoxa and Plausible Argumentation", *Argumentation*, vol. 12, no. 1, pp. 95–113.
- Walton, D., 2008, *Witness Testimony Evidence: Argumentation, Artificial Intelligence and Law*, Cambridge University Press, Cambridge, Mass.
- , 2006, *Fundamentals of Critical Argumentation*, Cambridge University Press, Cambridge, Mass.
- , 2002, *Legal Argumentation and Evidence*, Penn State Press, University Park, Pennsylvania.
- , 1998, *The New Dialectic: Conversational Contexts of Argument*, University of Toronto Press, Toronto.
- , 1992, "Rules for Plausible Reasoning", *Informal Logic*, vol. 14, no. 1, pp. 33–50.
- Walton, D., C.W. Tindale, y T.F. Gordon, 2014, "Applying Recent Argumentation Methods to Some Ancient Examples of Plausible Reasoning", *Argumentation*, vol. 28, no. 1, pp. 85–119.
- Walton, D. y C. Reed, 2008, "Evaluating Corroborative Evidence", *Argumentation*, vol. 22, no. 4, pp. 531–553.
- Wolpert, L., 1994, *The Unnatural Nature of Science*, Harvard University Press, Cambridge, Mass.
- Woods, J.H y D. Gabbay, 2005, *A Practical Logic of Cognitive Systems*, vol. 2, Elsevier, Amsterdam.
- Woods, J.H. y D. Walton, 1982, *Argument. The Logic of Fallacies*, McGraw Hill-Ryerson, Nueva York.

Recibido el 17 de enero de 2014; revisado el 27 de abril de 2014; aceptado el 13 de mayo de 2014.