
¿La Inferencia a la Mejor Explicación habilita desacuerdos racionales entre pares epistémicos?¹

Roberto Azar Romano / Universidad de Buenos Aires

Recibido el 20 de febrero de 2019. Aceptado el 17 de julio de 2019

> Resumen

Uno de los problemas de la epistemología social veritativista consiste en analizar si puede haber desacuerdos racionales (sin error) entre agentes que se consideran pares epistémicos sobre una cierta cuestión. Si un agente A_1 desacuerda con un agente A_2 al que considera igualmente inteligente, reconociendo que cuenta con la misma evidencia y con las mismas habilidades para evaluar correctamente dicha evidencia, se puede afirmar que hay una evidencia de segundo orden que lo conduce a pensar que él mismo puede estar equivocado. En filosofía de la ciencia, desde la segunda mitad del siglo XX, se ha incorporado en el debate que enfrenta a realistas y antirrealistas científicos el empleo de una regla conocida como “Inferencia a la Mejor Explicación” (IME). La IME, en su interpretación más directa, le exige al agente inferir la verdad (o al menos la verdad aproximada) de la mejor explicación disponible de la evidencia. Muchas versiones se han ofrecido de este mecanismo inferencial, desde la primera aplicación del concepto en Gilbert Harman (1965) hasta las más recientes elucidaciones por parte de autores como Peter Lipton (1991) y Alexander Bird (2007, 2010). El presente artículo pretende evaluar si son posibles los desacuerdos racionales entre pares epistémicos cuando se emplea con éxito una IME en el contexto científico.

» *Palabras clave: epistemología social veritativista, desacuerdos racionales, pares epistémicos, filosofía de la ciencia, inferencia a la mejor explicación.*

Does Inference to the Best Explanations enable rational disagreements between epistemic pairs?

> Abstract

One of the problems of the veritativist social epistemology is to analyze if there can be rational disagreements (without mistake) between agents who consider themselves epistemic peers on a certain issue. If an agent A_1 disagrees with an agent A_2 to whom he considers equally intelligent, recognizing that he has the same evidence and the same skills to correctly evaluate evidence, it can be said that there is a second order evidence that leads him to think that he may be wrong.

¹ Agradezco al Dr. Nicolás Lo Guercio (UBA), pues sus clases en el marco de un seminario de doctorado me han estimulado a investigar con mayor profundidad las temáticas que aborda el presente artículo.

In philosophy of science, since the second half of the twentieth century, the use of a rule known as “Inference to the Best Explanation” (IBE) has been incorporated into the debate that confronts scientific realists and anti-realists. IBE, in its most direct interpretation, requires the agent to infer the truth (or at least the approximate truth) of the best available explanation from the evidence. Many versions have been offered of this inferential mechanism, from the first application of the concept in Gilbert Harman (1965) to the most recent elucidations by authors such as Peter Lipton (1991) and Alexander Bird (2007, 2010). The present paper tries to evaluate if the rational disagreements between epistemic pairs are possible when an IBE is successfully used in the scientific context.

» *Keywords: veritativist social epistemology, rational disagreements, epistemic peers, philosophy of science, inference to the best explanation.*

> Introducción

Uno de los problemas de la Epistemología Social de corte veritativista consiste en analizar si puede haber desacuerdos racionales (sin error) entre agentes que se consideran pares epistémicos sobre una cierta cuestión. Si un agente A_1 desacuerda con un agente A_2 al que considera igualmente inteligente, reconociendo que cuenta con la misma evidencia y con las mismas habilidades para evaluar correctamente dicha evidencia, se puede afirmar que hay una *evidencia de segundo orden* que lo conduce a pensar que él mismo puede estar equivocado. Ante esta situación, hay un grupo de filósofos que argumentan que para conservar la racionalidad se hace menester revisar la creencia inicial (Cf. Feldman, 2006; Christensen, 2007), mientras que otro lote de pensadores asevera que se puede continuar en la racionalidad sin necesidad de revisar la creencia inicial, ya sea porque se apele a un conservadurismo epistémico (Harman, 1986) ya sea porque se invoque una suerte de egoísmo epistémico, de acuerdo con el cual las propias creencias gozan de un *plus epistémico* que las hace superiores para el agente que las porta (Bergmann, 2009; Huemer, 2011). Quienes se ubican en la primera línea argumentativa mencionada asumen una posición conciliacionista, según la cual una vez que A_1 y A_2 reconocen que han dado respuestas incompatibles ante la misma evidencia, tienen que modificar sus actitudes doxásticas iniciales para continuar siendo racionales.

En el campo de la filosofía de la ciencia, desde la segunda mitad del siglo XX, se ha incorporado en el debate que enfrenta a realistas y antirrealistas científicos², el empleo de una regla conocida como “Inferencia a la Mejor Explicación” (en adelante, IME). Como su nombre sugiere, la IME, en su interpretación más directa, le exige al agente inferir la verdad (o al menos la verdad aproximada) de la *mejor explicación* disponible de la evidencia. Muchas versiones se han ofrecido de este mecanismo inferencial, desde la primera aplicación del concepto en la pluma de Gilbert Harman (1965) hasta las más recientes elucidaciones por parte de autores de la envergadura de Peter Lipton (1991) y Alexander Bird (2007, 2010). El presente trabajo pretende evaluar si son posibles (o no) los desacuerdos racionales entre pares epistémicos cuando se emplea con éxito una IME en el contexto científico.

² No es el objetivo del presente trabajo profundizar en ese debate.

> Desacuerdos racionales entre pares epistémicos

Para comprender en qué consistirían los desacuerdos racionales sin error que nos interesan especialmente debemos pensar en una situación de paridad tanto evidencial cuanto cognitiva, es decir que no sólo se trata de que los dos agentes potenciales poseen exactamente la misma evidencia relevante, sino que, además, se saben portadores del mismo grado de inteligencia y con idénticas habilidades para evaluar aquella evidencia. El argumento, en resumen, se podría reconstruir de la siguiente manera:

- 1) A_1 y A_2 *desacuerdan* sobre alguna cuestión (evalúan la misma evidencia E , pero A_1 cree que p y A_2 cree que $\neg p$).
- 2) A_1 y A_2 son *pares epistémicos*.
- 3) A_1 y A_2 reconocen 1) y 2).

siendo A_1 y A_2 los agentes que evalúan la misma evidencia relevante, pero llegan a conclusiones incompatibles (p y $\neg p$).

Se asume en el debate que la justificación requerida debe entenderse aproximadamente como una racionalidad de tipo *epistémica*, siendo un supuesto de base que la meta epistémica por excelencia es “creer lo verdadero y evitar creer lo falso”. Por otra parte, se pueden distinguir dos sentidos de racionalidad atendiendo al factor temporal: una racionalidad *diacrónica* (a mediano o largo plazo) y una racionalidad *sincrónica* (simultánea). Es este último sentido del término ‘racionalidad’ el que se supone en el presente debate: una vez que A_1 (o A_2) obtuvo la evidencia de que el otro desacuerda con él y de que se trata de un par epistémico, ¿Cuál tiene que ser el impacto ahora? ¿Qué hacer? ¿Revisar las actitudes doxásticas iniciales o acaso mantenerlas guiado por argumentos independientes que vuelven más plausibles las propias creencias que las del oponente?

> Argumento conciliacionista

Como se adelantó, la posición conciliacionista (o conciliadora) postula la necesidad de modificar la actitud doxástica inicial una vez que ambos agentes caen en la cuenta de que dieron respuestas incompatibles ante la misma evidencia. Podríamos conjeturar que lo que ellos buscan es eliminar los desacuerdos racionales entre pares diacrónicamente, es decir, en el mediano y largo plazo. Pues es innegable que en el plano sincrónico los desacuerdos racionales entre pares acontecen. Y una vez que los agentes toman conciencia de ese hecho, los conciliacionistas argumentan, deben cambiar sus creencias iniciales para poder conciliar y evitar el incómodo desacuerdo. A continuación, explicitaremos los pasos de lo que podríamos llamar el ‘argumento conciliacionista’ (**AC**) y, luego, analizaremos qué supuestos se hallan implícitos en dicho argumento.

[AC]

1. A_1 y A_2 *desacuerdan*.
2. A_1 y A_2 son *pares epistémicos*.
3. 1. y 2. son reconocidos por A_1 y A_2 .
4. A_1/A_2 cree justificadamente que A_2/A_1 tiene la misma probabilidad de haber formado una creencia racional.
5. A_1 formó una actitud doxástica irracional³ o A_2 formó una actitud doxástica irracional.
6. Tengo una evidencia nueva (de segundo orden) de haber respondido incorrectamente a la evidencia inicial.

Ergo,

7. Debo conciliar.

Vale notar que [AC] no funciona sin al menos tres supuestos que se hallan a la base de los pasos 2, 4 y 5. A continuación, analizaremos cada uno de esos supuestos y comenzaremos a vislumbrar modos posibles de bloquear el argumento.

> Simetría

En primer lugar, queda claro que el paso 2. presupone un elemento básico que, en principio, debería aceptar cualquiera que esté en favor de la existencia de desacuerdos racionales entre pares epistémicos en el nivel sincrónico, a saber, que los agentes que participan en la discusión tienen una *simetría* absoluta que implica que (i) A_1 y A_2 disponen de la misma evidencia relevante; (ii) A_1 y A_2 tienen grados idénticos de inteligencia/competencia; (iii) A_1 y A_2 son igualmente buenos para evaluar la evidencia⁴; y (iv) las *actitudes doxásticas* (creencias y descreencias) de A_1 y A_2 poseen el mismo peso epistémico.

Se sigue que el primer recurso para bloquear [AC] podría consistir en negar la simetría entre A_1 y A_2 . Más adelante revisaremos algunas estrategias que argumentan en esa dirección.

³ En el sentido de no haber respondido correctamente a la evidencia.

⁴ Tal vez (ii) y (iii) puedan agruparse en una única característica, dado que, habiendo muchos tipos de inteligencia, cuando aquí se dice que ambos agentes son igualmente inteligentes, estamos circunscribiendo esta cualidad a las habilidades epistémicas que los sujetos poseen y que les garantiza, *prima facie*, evaluar correctamente la evidencia cuando se lo proponen.

> Tesis de independencia

El paso 4 da por sentado que vale la *tesis de independencia*, una de cuyas formulaciones posibles diría algo así como lo siguiente: “Cuando un agente A_1 está evaluando las credenciales epistémicas de su oponente A_2 , debe hacerlo sin fundarse para ello en el razonamiento que lo llevó a formar su creencia inicialmente”. De ahí que A_1 no pueda sino creer que su oponente tiene la misma probabilidad que él de haber formado una creencia racional. Afirmo que esta tesis va unida al principio de simetría expuesto anteriormente. En efecto, la mejor manera de negar la tesis de independencia consistiría en argumentar que hay casos en los que sí cabría confiar en nuestro propio razonamiento inicial y, consecuentemente, basarnos en él para evaluar las creencias portadas tanto por mí como por mi oponente potencial. Ahora bien, para poder afirmar fundadamente que mi razonamiento inicial tiene un valor epistémico *per se* debo negar la simetría absoluta entre los agentes que participan de la discusión. De modo que las estrategias que evaluaremos a continuación servirán para considerar intentos de refutar la tesis de independencia y, en el mismo acto, el principio de simetría.

> Tesis de unicidad

El paso 5 de [AC] consiste en una disyunción excluyente que presupone la que, siguiendo a Douven (2009), considero la pieza clave del argumento, a saber, la *tesis de unicidad* que se podría formular de la siguiente manera: “Dado cualquier cuerpo de evidencia y dada cualquier proposición, hay una ⁵ actitud doxástica que es racional adoptar”. Ciertamente se trata de una tesis tan fuerte como controvertida que resulta central tanto para comprender el argumento conciliacionista como para entender cuál es el significado exacto del término “IME”, del cual nos ocuparemos con un alto grado de detalle. Todo lo anterior justifica nuestro énfasis en analizar algunos modos posibles de refutar la tesis de unicidad. Argumentaremos en dos direcciones: desde el punto de vista teórico, afirmaremos con Douven (2009: 348) que sobre la base de nuestras mejores teorías actuales del apoyo evidencial –o teorías de la confirmación, como habitualmente se las llama– podemos decir que la tesis de unicidad es falsa. En segundo lugar, anclados en la práctica científica, sostendremos que, en múltiples ocasiones, los científicos, que cuentan con la misma evidencia relevante y con similares competencias epistémicas, no llegan a la conclusión de que haya una única actitud doxástica que sea racional adoptar, sino que suelen creer en más de una teoría igualmente compatible con la evidencia relevante y, contextualmente, deciden utilizar alternativamente una teoría u otra.

> Modos posibles de bloquear el argumento conciliacionista

⁵ Cabe notar que aquí hay una ambigüedad: mientras que Feldman (2007) formula la tesis sosteniendo que hay como *máximo* una actitud racional, White (2005) sostiene que hay exactamente una actitud racional. Estimo que la primera formulación es mejor que la segunda.

Hemos visto que las posturas conciliacionistas sostienen que un agente que pretende ser racional no puede apegarse a sus creencias iniciales una vez que sabe que desacuerda con un par epistémico, es decir, con una persona que cuenta con la misma evidencia relevante y con competencias epistémicas semejantes.

La reconstrucción de [AC] puso de manifiesto que quienes argumentan en esa dirección presuponen tres tesis cuanto menos controvertidas, a saber, el principio de simetría, la tesis de independencia y la tesis de unicidad. De donde se sigue que el ataque a cualquiera de ellas socavaría [AC] y, consecuentemente, daría lugar a la posibilidad de que los desacuerdos racionales se puedan afrontar sin necesidad de conciliar.

Hemos argumentado que el principio de simetría y la tesis de independencia están tan íntimamente vinculados que sólo funcionan puestos en conjunción. A continuación, en efecto, exponemos dos estrategias que atacan la tesis de independencia y, para hacerlo, niegan el principio de simetría. En primer lugar, consideremos el ‘conservadurismo epistémico’ defendido por Gilbert Harman (1986). Según esta posición, la estabilidad del propio sistema de creencias posee un valor epistémico *per se* que hace plausible no abandonar las creencias que se portan salvo que aparezcan alternativas con mayor apoyo evidencial. Dicho de otro modo, el método de formación de creencias propio se considera confiable. Tanto es así que el agente está racionalmente justificado a mantener su creencia inicial aun cuando se entere de que hay disponible una alternativa evidencialmente equivalente. Esta versión metodológica del ‘conservadurismo epistémico’ parece prometedora para pensar que hay auténticos desacuerdos racionales entre pares, dado que en ellos se cumple precisamente que ambos agentes poseen alternativas que resultan igualmente apoyadas por la evidencia relevante. Todo el punto, entonces, consiste en saber si el ‘conservadurismo epistémico metodológico’ es correcto o no lo es. El siguiente argumento sugiere la incorrección del principio considerado. Supongamos que dos sujetos, S_1 y S_2 , que cuentan con la misma evidencia relevante, poseen las siguientes actitudes doxásticas iniciales:

S_1 : p

S_2 : -p

Supongamos, además, que S_1 , motivado por la posición de Harman, decide mantener su creencia inicial (ante la presencia de una alternativa que resulta evidencialmente equivalente) en tanto que S_2 opta por modificar su creencia inicial y comienza a creer que p. En tal caso, ambos agentes creerían en la misma proposición (p). Pero, de acuerdo con el ‘conservadurismo epistémico metodológico’, sólo S_1 estaría racionalmente justificado en tanto que S_2 (aun poseyendo la misma evidencia que su oponente y creyendo exactamente lo mismo) no lo estaría por el solo hecho de haber cambiado su creencia inicial. Esto último resulta anti-intuitivo y muestra los límites de la postura conservadurista.

Una segunda estrategia que vía la negación de la simetría epistémica refutaría la tesis de independencia es la apelación a una suerte de ‘egoísmo epistémico’ (Bergmann, 2009; Huemer, 2011). En una de sus variantes (Cf. Bergmann, 2009), esta postura asume que un *insight* propio

posee un peso epistémico mayor por el solo hecho de ser propio y que no resulta compensado por el mero conocimiento de que el oponente posee un *insight* contrario al mío. La ‘asimetría epistémica’ postulada es evidente, pues por hipótesis no puede haber un peso epistémico equivalente entre un *insight* propio y la creencia de que el otro tiene un *insight*. Se podría objetar que este argumento contiene una petición de principio, puesto que, si S_1 y S_2 son internamente racionales, cabría suponer que lo que poseen no son dos *insights* contrarios, sino más bien dos juicios incompatibles acerca de qué debe contar como un *insight* genuino. En cualquier caso, una segunda variante de esta estrategia (Cf. Huemer, 2011) postula que hay algunas normas epistémicas que están centradas en el agente. Se trata de un ‘fenomenalismo conservativo’ de acuerdo con el cual la fuente básica de evidencia con la que contamos son los *seemings*⁶. Como consecuencia, la norma epistémica favorecida diría algo así como lo siguiente: “si al agente le parece que p , tiene razones *prima facie* para creer que p ”. Una vez más, se niega la asimetría epistémica entre los agentes involucrados en un eventual desacuerdo, pero esta vez por medio de la postulación de una relación especial o privilegiada entre el agente (considerado en primera persona) y sus propios *seemings*. La ‘asimetría epistémica’, entonces, descansaría en una norma egoísta que autoriza al agente a otorgarle mayor peso epistémico a sus propios *seemings* que a los de su oponente. Sin dudas, la aceptación de esta norma epistémica egoísta niega automáticamente la tesis de independencia, dado que ya no se sostiene la creencia de que mi rival tenga la misma probabilidad que yo de haber formado una creencia racional. Mi especial relación con mis *seemings*, por el contrario, me conducirá a creer que mi probabilidad de haber formado una creencia racional es mayor. El problema es que mi rival pensará lo inverso (que su probabilidad de haber formado una creencia racional es mayor que la mía). De donde se sigue que no habría medios objetivos para determinar quién está más cerca de la verdad en un eventual desacuerdo.

Un contraejemplo, a mi juicio, más eficaz que las estrategias anteriormente consideradas para refutar la tesis de independencia no se centra en la confiabilidad del método empleado, sino más bien en la *cantidad de desacuerdos* entre los sujetos involucrados en la discusión. Según este contraejemplo, si chequeamos que nuestro desacuerdo con otro agente no se limita al caso que se está considerando, sino que los desacuerdos se produjeron una cantidad reiterada de veces, la intuición más apropiada me lleva a suponer que mi rival no era en realidad un par, sino que yo estaba equivocado al considerarlo como tal.

Cabe destacar que la tesis de independencia es una condición necesaria para que [AC] funcione, pero no es una condición suficiente. En efecto, no nos garantiza que tengamos que conciliar. Sostenemos que la pieza clave de [AC] es la tesis de unicidad, lo cual justifica que le dediquemos buena parte de este trabajo a evaluar sus credenciales. Como se adelantó, argumentaremos en dos direcciones: en primer lugar, sostendremos con Douven (2009: 348) que nuestras mejores teorías actuales del apoyo evidencial sugieren que la tesis de unicidad es falsa. En segundo lugar, analizaremos qué es lo que acontece en la práctica científica cuando dos (o más) científicos con similares competencias epistémicas se enfrentan con la misma evidencia relevante. En definitiva, cuestionaremos que la tesis de unicidad (interpretada en un sentido fuerte) valga en la actividad científica real.

⁶ Se podría traducir esta expresión inglesa al castellano como “pareceres”.

Se supone que las teorías de la confirmación especifican, para cualquier pieza de evidencia E , y para cualquier hipótesis o creencia H , si E apoya (o confirma) H , es decir, si E es o no una evidencia para H . Una teoría de la confirmación podría, además, especificar si E apoya H fuertemente/moderadamente/débilmente, si E es para H una evidencia más fuerte que E' , o si E es una evidencia más fuerte para H que para H' . También podría hacer cuantitativamente precisa la medida en la cual E apoya H .

Durante mucho tiempo se ha pensado que para que una teoría de la confirmación fuera adecuada debiera ser completamente *objetiva* en el sentido de que muestre que la evidencia para una creencia es independiente de quien “usa” una teoría. En las décadas de 1940 y 1950, tanto Carnap como Hempel, los filósofos de la ciencia más eminentes de ese tiempo, dedicaron gran parte de su trabajo al desarrollo de una teoría de la confirmación objetiva. Así, Carnap (1950) propuso una noción de ‘probabilidad lógica’ entendida como un tipo de probabilidad que estaba –supuestamente– determinada, para cualquier pieza de evidencia y para cualquier hipótesis dada, sobre la base *únicamente* de las relaciones lógicas existentes entre los enunciados que expresan la evidencia y los que expresan las hipótesis. De la misma manera, Hempel (1945) intentó caracterizar la relación de confirmación en términos estrictamente lógicos. Sin embargo, a estas concepciones no les fue mejor que al antiguo método hipotético-deductivo (como se defiende en Ayer (1936), por ejemplo) al que debían sustituir, una teoría que también proporcionó una definición lógica de la ‘confirmación’, pero a la que generalmente se había considerado como excesivamente simplista. El hecho de que ese modelo fracasara al aplicarse a las llamadas ‘hipótesis teóricas’, como así también la ‘paradoja del cuervo’ (*raven paradox*⁷), son solo dos de los problemas más conocidos que resultaron fatales para la concepción de Hempel. Y Carnap se vio incapaz de determinar de una manera no arbitraria las probabilidades lógicas supuestamente únicas incluso para lenguajes muy simples; de hecho, ni siquiera fue capaz de descalificar asignaciones de probabilidad ostensiblemente absurdas. En respuesta a estos fracasos, los filósofos de la ciencia se hicieron cada vez más empáticos con las teorías de la confirmación que dejaban cierto margen para los elementos subjetivos. Tanto es así que no resulta exagerado afirmar que las teorías de la confirmación actualmente más populares son todas de este tipo. El ejemplo presumiblemente más claro a este respecto es la teoría de la confirmación bayesiana. Según ella, E es una evidencia para H si y sólo si la probabilidad de H condicionalizada sobre E es mayor que la probabilidad de H sin condicionalizar, lo que a menudo se escribe como $\Pr(H | E) > \Pr(H)$; E es una evidencia en contra de H , o E disconfirma H , si y sólo si $\Pr(H | E) < \Pr(H)$; y E es evidentemente irrelevante para H de otro modo. Esto todavía podría definir una noción objetiva de apoyo evidencial, si no fuera por el hecho de que la teoría toma a \Pr como representando una función de probabilidad subjetiva del agente, es decir, su función de grados de creencia racional. Desde la perspectiva bayesiana, en definitiva, la relación de apoyo evidencial es subjetiva: la evidencia apoya (o fracasa en apoyar) creencias siempre relativamente a la función de grados de creencia de un agente, donde no existe una única función correcta. Un tipo diferente de ejemplos de teorías de la confirmación subjetivas proviene de todas las teorías que comprenden, o se basan en, la

⁷ Esta famosa paradoja está conectada con el problema de la confirmación de los enunciados universales por un enunciado singular. Así, por ejemplo, las consideraciones puramente lógicas determinarían que la observación de que “hay zapatos blancos” confirma la hipótesis de que “Todos los objetos no negros son no cuervos” y, así, por equivalencia “Todos los cuervos son negros”. Esta perspectiva, que sugiere que la observación de “zapatos blancos” confirma la hipótesis de que “todos los cuervos son negros” es considerada inconsistente con el concepto de confirmación.

regla de inferencia generalmente llamada “Inferencia a la mejor explicación” (IME). Si bien nos ocuparemos de sus distintas formulaciones luego, la interpretación más directa de la IME nos prescribe inferir la verdad (o la verdad aproximada) de la mejor explicación de la evidencia disponible. Por eso Douven piensa que la IME se podría considerar como un genuino principio teórico de confirmación, ya que es viable pensarla como una definición implícita de la relación de apoyo evidencial en términos de explicación, más o menos de la siguiente manera: *E* apoya *H* si y sólo si *H* es la que mejor explica *E* (Douven, 2009: 350). Para ver por qué esto podría hacer que la relación de apoyo evidencial sea subjetiva, debemos notar que la aplicación de la IME exige manifiestamente que clasifiquemos todas las explicaciones disponibles de acuerdo con su bondad (*goodness*). Al menos se supone que debemos juzgar cuál de ellas es la mejor en cuanto a explicar la evidencia; de ahí que la seleccionada como ‘mejor explicación’ quedaría mejor respaldada por dicha evidencia. Ahora bien, es una cuestión controvertida, y ampliamente debatida aun hoy en día, cuál es la noción de ‘explicación’ subyacente a la aplicación de la IME. Pero desde una concepción sobre la explicación bastante popular, la *bondad explicativa* arriba mencionada consiste en satisfacer las llamadas ‘virtudes teóricas’, como la simplicidad, la fecundidad, la abarcabilidad, la coherencia tanto con el conocimiento de fondo como así también con las suposiciones metafísicas de base, y, por supuesto, la consistencia con la evidencia a ser explicada. El punto problemático es que esas virtudes son en cierta medida subjetivas; no existen estándares totalmente objetivos de simplicidad, fecundidad, etcétera, y por lo tanto cada individuo puede interpretarlas de un modo un poco diferente. Además, no existe una forma objetivamente correcta de ponderar estas virtudes entre sí en los casos en los que apuntan en direcciones diferentes. Como resultado, no existen criterios completamente objetivos para juzgar la ‘bondad explicativa’ (Douven, *Ibid.*). Como veremos en ocasión de presentar la caracterización de la IME debida a Lipton (1991), una concepción más reciente acerca de la ‘explicación’ enfatiza la relación entre la explicación y la comprensión. Desde este punto de vista, entonces, la mejor explicación sería aquella que proporciona la mayor comprensión potencial de la evidencia. Sin embargo, salvo que se elucide en términos completamente objetivos la noción de ‘comprensión potencial’ –cosa que hasta el momento no se ha logrado–, el concepto de ‘comprensión’ es *prima facie* de índole subjetivo. Y no se ve con claridad qué nos garantizaría que lo que hace que la evidencia sea más comprensible para usted sea también lo que hace que sea más comprensible para mí.

Lo dicho hasta aquí sugiere que la Unicidad (al menos en una versión fuerte) es insostenible, y que es posible un desacuerdo racional si los factores subjetivos subyacen a las relaciones de apoyo evidencial de la forma en que lo hacen, de acuerdo con las teorías de la confirmación antes mencionadas. Pues si bien la unicidad no es en sí misma una tesis sobre el apoyo evidencial, sería absurdo sostener que ningún cuerpo de evidencia *E* permite más de una actitud racional hacia una proposición *H* aunque la pregunta de si *E* apoya a *H* no tiene, en general, una única respuesta correcta (Douven, *Ibid.*: 351).

Antes de dirigirnos al contexto de la práctica científica, advertamos que la IME debería poder seleccionar (preferentemente) *una hipótesis* a partir de muchas explicaciones posibles satisfactorias. Si esto fuera así, podríamos verlo como una aplicación particular de la tesis de unicidad. Ahora bien, ¿habrá un mecanismo de selección que nos permita elegir unívocamente

una hipótesis explicativa a partir de un conjunto dado o habrá condiciones contextuales que influirán directamente en dicha selección pudiendo dar lugar a resultados diferentes? Tomemos el siguiente ejemplo de Johannes Persson (2007: 142). En un contexto dado, podríamos contar con una explicación funcional y con una explicación físico-química sobre el funcionamiento de un termostato:

Algunos filósofos dirían que la explicación mecánica es la mejor, y que la explicación funcional no es tan buena. Wright sostiene que la explicación funcional no es necesariamente reemplazada por la explicación mecánica; ambas son legítimas y podrían complementarse. En un contexto una de ellas podría ser mejor; en otro contexto la otra podría ser preferible. La frase 'inferencia a la mejor explicación' involucra una tesis de unicidad que es difícil de justificar (Salmon, 2001: 67, citado por Persson, la traducción es mía)

Cuando en el texto citado se habla de una 'tesis de unicidad' (*uniqueness claim*), se quiere significar que la expresión misma 'inferencia a la mejor explicación' parece conducir a la idea según la cual el proceso inferencial debe llevar al resultado de la selección de *una única explicación* como *la* explicación adecuada. Pero en el ejemplo considerado, habida cuenta de que ambas explicaciones son compatibles, no se ve con claridad por qué deberíamos impedir la inferencia hacia ambas. Si, por otra parte, tomamos en consideración el lema de Lipton (uno de los grandes referentes en los estudios sobre la IME), según el cual "La expresión 'Inferencia a la Mejor Explicación' debe ser sustituida por la frase más apropiada, aunque menos memorable, 'inferencia a la mejor de las explicaciones en competencia disponibles *cuando la mejor es suficientemente buena*'⁸ (Lipton, 2004: 104, el subrayado es mío), la tesis de unicidad se derrumba (al menos para el caso considerado y para otros casos análogos).

Por lo demás, este punto también fue anticipado por Thomas Kuhn (1977). Debido a que él creía que los factores subjetivos eran esenciales para la confirmación, una de las conclusiones a las que arribó es que científicos igualmente racionales que tienen acceso a la misma evidencia podrían, sin embargo, llegar a conclusiones opuestas sobre una cuestión específica:

Cuando los científicos deben elegir entre teorías en competencia, dos hombres totalmente comprometidos con la misma lista de criterios de elección [como la simplicidad, la fecundidad y la consistencia] pueden, sin embargo, llegar a conclusiones diferentes. Quizás ellos interpreten la simplicidad de manera diferente o tengan diversas convicciones sobre el rango de campos dentro del cual debe cumplirse el criterio de consistencia. O tal vez estén de acuerdo sobre estas cuestiones, pero difieran sobre los pesos relativos que se otorgarán a estos u otros criterios cuando se utilicen conjuntamente. (Kuhn, 1977: 324, citado por Douven, la traducción es mía)

Y de una manera aún más reveladora, más adelante en el mismo texto Kuhn argumenta que la consideración del desarrollo de teorías recientemente sugeridas en la ciencia muestra que antes de que la comunidad científica aceptara una nueva teoría

[dicha teoría] ha sido probada en el tiempo por la investigación de un número de hombres, algunos

⁸ Volveremos sobre esta cuestión más adelante.

trabajando en ella, otros dentro de su rival tradicional, y ese modo de desarrollo requiere un proceso de decisión que permita a los hombres racionales estar en desacuerdo, y tal desacuerdo sería excluido por el algoritmo compartido [es decir, la teoría de la confirmación objetiva] que los filósofos generalmente han buscado” (Kuhn, 1977: 332).

Como se ve, tanto las teorías de la confirmación más populares como así también la práctica científica real parecen testimoniar en contra de la tesis de unicidad. Ahora bien, ¿dicha tesis estará efectivamente presupuesta en el empleo de la IME? Si esto fuera así, ¿Qué impacto tendría en relación con la posibilidad de que haya desacuerdos racionales entre pares epistémicos en casos en los que la IME se aplica de manera exitosa? Para responder este tipo de interrogantes, debemos profundizar en las diferentes versiones que sobre la IME se han presentado históricamente.

> Versiones de la ‘Inferencia a la Mejor Explicación’ (IME)

Origen del concepto en la filosofía de Gilbert Harman [1965]

Gilbert Harman empleó por primera vez el término “IME” en su artículo *The Inference to the Best Explanation* (1965). Su objetivo explícito en ese entonces era exponer su posición sobre la inducción enumerativa⁹. Más específicamente, él defendió dos tesis íntimamente vinculadas entre sí: 1. Las inferencias que parecen ser instancias de la inducción enumerativa serían mucho mejor explicadas como inferencias a la mejor explicación, y 2. En la selección de una hipótesis que explique cierta evidencia frecuentemente hacemos uso de determinados lemas¹⁰, que sólo se ponen de manifiesto si consideramos el proceso de selección de hipótesis como una IME. O sea que interpretar tales inferencias como meras inducciones enumerativas enmascara el importante rol que cumplen los lemas para que arribemos a una explicación que constituya un auténtico conocimiento. Para clarificar esto, podemos pensar en un ejemplo simple de la vida corriente: supongamos que advertimos que un amigo está empapado y explicamos la evidencia diciendo que está lloviendo en este momento. Si hubiera aplicado tan solo una inducción típica, estableciendo un examen de las correlaciones pasadas (*i.e.*, usualmente alguien está mojado porque está lloviendo sobre esa persona), entonces habría oscurecido el rol de lemas muy importantes: que mi amigo se encontraba al aire libre y que no llevaba un paraguas ni una campera de lluvia. Es evidente que esos lemas intermedios –así los llama Harman– juegan un papel relevante en mi explicación, pues yo no habría postulado la lluvia como causa de la humedad de mi amigo si hubiera pensado que él estaba equipado con un buen paraguas. Otro tipo de ejemplos en el mismo sentido proviene de lo que Harman llama la creencia *por autoridad*¹¹. La autoridad en cuestión podría ser o bien una persona que es experta en su campo de estudio o bien un libro de referencia en alguna disciplina. Es indudable que gran parte del conocimiento que poseemos está basado en autoridades en este sentido y en los casos en los cuales nos confiamos del

⁹ Cabe aclarar que Harman habla de “inducción enumerativa” toda vez que se infiere de la regularidad observada la regularidad universal o al menos la regularidad en el próximo caso. Actualmente, con un lenguaje más preciso, diríamos que ciertas inferencias que él etiqueta como inducciones enumerativas en realidad deberían ser vistas como inducciones por analogía.

¹⁰ En el contexto argumentativo de Harman podríamos definir un *lema* como “un enunciado filosófico que se acepta como verdadero para averiguar si otro enunciado es verdadero”. El concepto quedará claro cuando analicemos los ejemplos brindados por el propio autor para iluminar su empleo.

¹¹ Cabe aclarar que este tipo de conocimiento no sólo se adquiere en la vida cotidiana, sino también en la vida académica.

conocimiento así obtenido no se nos puede acusar de cometer la famosa falacia no formal *ad verecundiam*¹². Cuando un experto nos dice algo acerca de cierta materia, o cuando leemos algo en un libro de referencia en un área determinada de estudio, frecuentemente estamos justificados para creer que lo que se nos dice o lo que leemos es correcto. Que nuestra creencia sea verdadera es la primera condición que se tiene que cumplir para que nuestro conocimiento sea auténtico. Una segunda condición establece que lo que se nos dice o lo que leemos no debe estar allí por error. Es decir, el hablante no debe haber tenido un lapsus que afecte el sentido de lo que dijo ni nuestra creencia puede estar basada en la lectura de una errata en un libro de texto. Aun si el lapsus o la errata han mudado una falsedad en una verdad, por accidente, no podemos afirmar razonablemente que tenemos conocimiento. Esto indica, según Harman, que la inferencia que hacemos desde el testimonio hacia la verdad debe contener como lema la proposición de que “la emisión está allí porque es creída y no a causa de un lapsus o de un error de imprenta” (Harman, 1965: 93).

La IME puede describirse como un proceso inferencial que comienza con la existencia de varias hipótesis en competencia $-H_1, H_2, \dots, H_n-$, todas las cuales pretenden explicar la misma evidencia intrigante $-E-$. Luego, a través de la evaluación de las virtudes explicativas de las hipótesis en competencia, se procede al rechazo de todas las hipótesis alternativas y se culmina (en el caso ideal) con la selección de la mejor hipótesis explicativa, es decir, aquella que proporciona la mejor explicación de la evidencia¹³. El concepto de “mejor explicación” sugiere que H es verdadera o al menos que tenemos buenas razones para creer que H es verdadera. Harman lo expresa en los siguientes términos: “uno infiere, a partir de la premisa de que una hipótesis dada proporcionaría una *mejor* explicación para la evidencia que cualquier otra hipótesis, la conclusión de que la hipótesis es verdadera” (Harman, 1965: 89). En general, agrega el autor, habrá muchas hipótesis que podrán explicar la evidencia, así que uno debe ser capaz de rechazar todas las hipótesis alternativas antes de estar autorizado a realizar la inferencia. Además,

La ‘inferencia a la mejor explicación’ corresponde *aproximadamente* a lo que otros han llamado ‘abducción’, ‘método de hipótesis’, ‘inferencia hipotética’, ‘método de eliminación’, ‘inducción eliminativa’ e ‘inferencia teórica’. Yo prefiero mi propia terminología porque creo que evita la mayoría de las connotaciones engañosas de las terminologías alternativas (Harman, *Ibid.*: 88, el subrayado es mío)

La cita evidencia que, a pesar de que muchos autores se empeñaron en identificar la IME con la abducción, el mismo Harman se cuidó de no plantear semejante identificación¹⁴. En efecto, él inventa un término absolutamente novedoso en su tiempo para dar cuenta de un proceso inferencial cuyo rasgo definitorio viene dado por la función selectiva. Harman reconoce que hay un problema en cómo definir “mejor”. Si bien él solo trata la cuestión al pasar, mencionando unos pocos criterios de selección (simplicidad, plausibilidad intrínseca, carencia de carácter *ad hoc*, etcétera), adelanta un gran problema que Lipton va a retomar más adelante.

12 En efecto, el carácter falaz de un razonamiento es contextual. Y si bien es claro que apelar a la autoridad de Einstein, pongamos por caso, para arribar a conclusiones sobre Política o Economía resultaría falaz, no lo sería acudir a la palabra del mencionado científico célebre si lo que queremos es resolver un problema que se presenta en el campo de la física relativista.

13 Por eso muchos sostienen que la IME puede entenderse como una regla que afirma aproximadamente lo siguiente: dada la evidencia intrigante (E) y dado un conjunto de hipótesis potenciales H_1, H_2, \dots, H_n que explican E, si H8 es una mejor explicación de E que cualquiera de las otras hipótesis, infiérase que H8 está más cerca de la verdad que cualquiera de esas otras hipótesis.

14 Para el lector interesado en las relaciones existentes entre los conceptos de ‘IME’ y ‘Abducción’, ver Azar, R. M. (2017).

Una nueva visión de la IME: Peter Lipton [1991]

La mayoría de los especialistas en el área coinciden en afirmar que para todos aquellos que trabajan en torno a la elucidación del concepto de “inferencia a la mejor explicación”, el texto *Inference to the best explanation* (1991) de Peter Lipton organiza la agenda. Su caracterización oficial de la IME establece lo siguiente:

Dados nuestros datos y nuestras creencias de fondo, inferimos la explicación que, si fuera verdadera, proporcionaría *la mejor* de las explicaciones en competencia que podamos generar de esos datos (siempre y cuando la mejor sea suficientemente buena *para nosotros* como para realizar alguna inferencia) (Lipton, 1991: 58, el subrayado es mío)

La pregunta que surge inmediatamente es la siguiente: ¿cómo seleccionamos entre las explicaciones potenciales? Según Lipton, la IME es un proceso que consta de dos etapas y ambas constituyen auténticos filtros de explicaciones potenciales:

En la primera etapa la capacidad imaginativa de los científicos generaría todas las explicaciones potenciales *plausibles*¹⁵, de donde se sigue que se filtrarían las explicaciones potenciales implausibles.

En la segunda etapa, los científicos analizarían las explicaciones potenciales plausibles que ya han pasado exitosamente el primer filtro y, finalmente, las clasificarían de acuerdo con su bondad explicativa en vistas de seleccionar la mejor explicación como la explicación verdadera (o, al menos, aproximadamente verdadera).

¿En qué consiste la bondad explicativa mencionada en la segunda fase? Lipton la caracteriza en términos de *loveliness*¹⁶. La mejor explicación es la explicación más adorable/más encantadora (*the loveliest explanation*). Tal vez resulte sorprendente que el autor utilice una expresión que parece responder más a cuestiones estéticas que epistemológicas. Pero, en cualquier caso, cabe preguntarse a qué se refiere exactamente con esa propiedad. Lipton afirma que se trata de una virtud explicativa que podríamos caracterizar de la siguiente manera: la *loveliness* es el atributo que posee la explicación que nos proporciona el mayor grado posible de comprensión de la evidencia.

En el texto citado más arriba, por otra parte, se subrayó deliberadamente la expresión ‘para nosotros’ en vistas de enfatizar el carácter aparentemente subjetivo que comporta lo que podríamos llamar el ‘requisito de satisfactoriedad’, de acuerdo con el cual “la explicación seleccionada como la mejor (a partir de un lote previamente constituido) debe ser lo suficientemente satisfactoria para nosotros como para que consideremos que vale la pena efectuar la inferencia”. De esto se sigue que a pesar de lo atractivo que resulta el lema “Inferencia a la mejor explicación”, en ciertos contextos puede ocurrir que ninguna de las explicaciones efectivamente disponibles

¹⁵ Al menos todas aquellas que los científicos sean capaces de imaginar.

¹⁶ Dada la dificultad para encontrar una traducción precisa en español, en general utilizaré la expresión “*the loveliest explanation*” o “*loveliness*” (para referirme a la propiedad que Lipton le atribuye a la mejor explicación) sin traducción alguna.

resulte favorecida¹⁷. Pero ¿de qué dependerá el hecho de que una explicación determinada pueda cumplir el ‘requisito de satisfactoriedad’? Mucho se ha insistido en ciertas virtudes explicativas que las hipótesis deberían poseer: simplicidad, coherencia con el conocimiento de fondo, poder predictivo, etcétera. Sin embargo, en la caracterización de Lipton la “*loveliness*” se perfila como la virtud explicativa privilegiada. El problema que surge es el siguiente: ¿Cómo saber que la explicación que goza del grado más alto de *loveliness* es también la más probablemente verdadera? Este interrogante pone de manifiesto un aspecto de su posición que resulta tan interesante como controvertido, a saber, la distinción que el autor realiza entre *likeliness* y *loveliness*. En palabras de Lipton

Los criterios de *likeliness* y *loveliness* podrían seleccionar la misma explicación en una competición particular, pero ellos son claramente diferentes tipos de estándares. La *likeliness* habla de la verdad; la *loveliness* de la comprensión potencial. Además, los criterios frecuentemente seleccionan explicaciones diferentes (Lipton, 1991: 59).

Si bien muchos podrían pensar *a priori* que la mejor explicación debe identificarse siempre con la más probable, Lipton aclara que, en ocasiones, la explicación más probable es poco esclarecedora. Así, por ejemplo, es muy probable que fumar opio haga que las personas duerman debido a sus poderes dormitivos, pero se trata de una explicación muy poco encantadora (*unlovely*). Inversamente, podría ocurrir que la explicación que goce del grado más elevado de *loveliness* sea muy poco probable. Quizás algunas teorías conspirativas proporcionen ejemplos de esto (Cf. Lipton, *o.c.*: 60).

Una de las razones por las cuales la probabilidad (*likeliness*) y el encanto explicativo (*loveliness*) frecuentemente divergen es que mientras que la *likeliness* es relativa a la evidencia total disponible, la *loveliness* no lo es, o al menos no en el mismo sentido. De modo que una explicación que posea al mismo tiempo los dos atributos mencionados en un momento t_0 , podría volverse poco probable en un momento t_1 en el cual evidencia adicional se ha descubierto. Piénsese en la mecánica newtoniana y el reciente advenimiento de la relatividad especial con los nuevos datos empíricos que la apoyan. La nueva evidencia disponible hace *menos probable* a la mecánica newtoniana en relación con la física relativista. Sin embargo, la explicación newtoniana continúa siendo igualmente encantadora, dado que su extraordinario “poder unificador” es tal que permite explicar no sólo cómo se producen las mareas, sino también cómo se mueven los planetas y cómo es la caída de los cuerpos en la tierra. De modo que, siguiendo cada una de estas dos cualidades puede, eventualmente, seleccionarse a la misma hipótesis como la mejor explicación: podría ser la más verosímil y también la más encantadora. Pero también puede no ser así, como en el caso mencionado. Hasta tanto no se logre proporcionar una caracterización objetiva del concepto de “*loveliness*”, creo que el problema permanecerá irresoluble en el marco de esta posición.

¹⁷ Esto atendería al menos contra una de las formulaciones posibles de la tesis de unicidad, a saber, aquella que postula que “dada cualquier evidencia y dada cualquier proposición, hay exactamente una actitud doxástica racional” (White, 2005)

Inferencia a la Única Explicación [Alexander Bird (2007, 2010)]

Alexander Bird (2007, 2010), quien describe a la IME de Peter Lipton como un paradigma¹⁸ en la filosofía de la ciencia, presenta un caso especial de la IME que, según él, nos permitiría seleccionar solo una explicación potencial, dado que sería *la única* explicación potencial auténticamente consistente con la evidencia. Se trata de la Inferencia a la única explicación (IUE), a la que también denomina “inferencia holmesiana”. De modo que Bird propone complementar la IME con un razonamiento eliminativo que sustituye el criterio de potencial explicativo por una inferencia denominada “holmesiana”. La idea se inspira en el lema de Sherlock Holmes “Elimina lo imposible y, lo que queda, aunque improbable, debe ser verdadero” (Conan Doyle, 1953: 94-118). Pero, ¿es esto posible en alguna circunstancia? El autor presenta dos casos históricos tomados de la historia de la medicina que, según él, avalan la tesis de que la IUE tiene lugar en la práctica científica, a saber, el famoso caso del Dr. Semmelweis en el hospital general de Viena (entre 1844 y 1848) y, más recientemente, el descubrimiento del HIV (1986).

La estructura básica de una inferencia holmesiana, según Bird, es la siguiente:

- (i) El hecho sorprendente E tiene una explicación (Determinismo)
- (ii) H_1, \dots, H_n son las únicas hipótesis que podrían explicar E.
- (iii) H_1, \dots, H_{n-1} han sido refutadas/falsadas por la evidencia (Falsación)

Por lo tanto

- (iv) H_n explica E.

La premisa (i) asume que el hecho intrigante o sorprendente (E) tiene una explicación. Es decir, quien aplica este proceso inferencial parte del presupuesto metafísico determinista, que constituye la negación de lo que podríamos llamar, siguiendo a Bird, la “hipótesis nula” (H_0). H_0 establecería que no hay ninguna explicación del fenómeno en cuestión. Dicho de otro modo, el determinismo excluye la posibilidad de que se trate de una casualidad. La tesis determinista parece funcionar en el plano de los eventos macroscópicos. Sin embargo, cuando nos trasladamos al nivel de los eventos atómicos o subatómicos, advertimos que no todos los hechos poseen una explicación, dado que la ocurrencia de ciertos eventos microscópicos es intrínsecamente indeterminista; tan sólo ocurre. Esto se pone de manifiesto, por ejemplo, al considerar el caso de la descomposición de un núcleo fisible (Bird, 2010: 350).

La premisa (ii) correspondería a la fase inicial de la IME liptoniana, dado que hace referencia a la conformación del conjunto de hipótesis explicativas plausibles que dio como resultado el primer filtro racional llevado a cabo por los investigadores. Este paso constituye una condición necesaria, aunque no suficiente, para que el proceso inferencial pueda completarse.

¹⁸ En este contexto sería mucho más atinado entender el término “paradigma” no en el sentido en el que lo emplea Thomas Kuhn, sino remitiéndonos al origen etimológico de la palabra *παράδειγμα*: *parádeigma* en griego antiguo significa “modelo” o “ejemplo”.

Sin dudas, la premisa (iii) es la específica de la inferencia holmesiana. Y supone un hecho difícilmente sostenible (al menos para los defensores de la tesis Duhem-Quine): que la evidencia permite ir refutando progresivamente todas las hipótesis rivales, salvo una, *la explicación verdadera*.

Si bien es altamente dudoso que se pueda demostrar que las premisas (i), (ii) y (iii) sean verdaderas, es claro que, si lo fueran, entonces la conclusión sería necesariamente verdadera. En efecto, la inferencia holmesiana es deductiva, razón por la cual Conan Doyle describe a su héroe como un maestro de la *deducción* (Bird, 2010: 350).

> ¿Son posibles los ‘desacuerdos racionales entre pares’ (DREP) cuando se aplica una IME de manera exitosa?

¿La IME presupone ‘Unicidad’?

Cabe pensar que la respuesta a este interrogante depende de la versión de la IME que se considere como así también de la definición de ‘Unicidad’ que se asuma. Ya hemos visto que cierta variante fuerte de la unicidad es incompatible con lo que acontece en la práctica científica real. También hemos adelantado en una nota al pie que, de acuerdo con la caracterización de White (2005) la Unicidad sostiene que “dado un cuerpo de evidencia y dada cualquier proposición, *hay exactamente una actitud racional* que es posible asumir”. Si supusiéramos esta versión (fuerte) de la Unicidad, creo que podríamos concluir rápidamente que la única versión de la IME que claramente presupone Unicidad es la que sostiene Bird (la IUE¹⁹), en tanto que las más débiles caracterizaciones de Harman y de Lipton permiten *prima facie* que en circunstancias en las cuales no haya ninguna explicación *lo suficientemente buena*, los investigadores suspendan el juicio. E incluso la perspectiva de Lipton admite que, en casos en los que haya más de una explicación adecuada en competencia –como sucedió con la explicación mecanicista y la explicación funcional–, se pueda proceder a la inferencia hacia ambas.

Sin embargo, estimo más apropiada la caracterización más débil de la ‘Unicidad’ debida a Feldman (2007). De acuerdo con esta, “dado cualquier cuerpo de evidencia y dada cualquier proposición, *hay como máximo una actitud doxástica* que es racional adoptar”. Esta definición de ‘unicidad’ en principio resulta incompatible con la IME de Lipton, dado que ya hemos visto que hay circunstancias en las cuales más de una actitud doxástica racional es pasible de ser asumida. Cabe recordar, no obstante, que Lipton también afirma explícitamente que el lema ‘Inferencia a la Mejor Explicación’ debería modificarse por la máxima más apropiada (aunque menos atractiva) “inferencia a la mejor de las explicaciones en competencia *cuando la mejor es suficientemente buena*”. En definitiva, lo que el autor está evidenciando es el hecho de que muchas veces ninguna explicación merece ser favorecida. Esto nos evoca uno de los problemas más serios que se le presentan históricamente a la IME: dado que, en la mayoría de los casos, no tenemos cómo considerar todas las posibles hipótesis que explicarían la evidencia –ellas serían probablemente

¹⁹ Recordemos que la IUE sólo funciona si al final del proceso hay solamente una hipótesis explicativa consistente con la evidencia por haber resistido con éxito todos los intentos de refutación.

infinitas–, ¿cómo saber si la mejor explicación está dentro de aquellas que hemos considerado? ¿y si la elegida como la mejor fuese la mejor dentro de un mal lote? Le debemos el <<argumento del mal lote>> a Bas van Fraassen. En palabras del autor:

Permítasenos conceder que los científicos han efectuado un ordenamiento de un conjunto de teorías T_1, \dots, T_n , todas las cuales ofrecen explicaciones potenciales de la evidencia e y que ellos han determinado cuál es la mejor explicación de e , digamos T_1 . Para que ellos puedan decir que T_1 es la explicación aproximadamente verdadera de e , ellos deben dar ‘un paso más allá del juicio comparativo de que T_1 es mejor que sus rivales reales’. Este paso involucra la creencia de que es más probable que la verdad se encuentre dentro del lote de teorías disponible para ellos. Pero nuestra mejor teoría bien puede ser ‘la mejor de un mal lote’. Así, para que el defensor de la IME afirme que la IME conduce a la verdad, él debe asumir un Principio de Privilegio. Es decir, debe suponer que ‘la naturaleza nos predispone a dar con el rango correcto de hipótesis’ (van Fraassen, 1989: 142, la traducción es mía).

No caben dudas de que se trata de un desafío muy serio para todo aquel que desee defender la idea de que la IME es un modo de inferencia que conduce a resultados confiables, al menos si se persiste en el empeño de identificar “mejor explicación” con “explicación verdadera”. ¿No radicará aquí el verdadero problema? Un espíritu realista científico tiende a portar una excesiva confianza en la vinculación existente entre la seleccionada como la mejor hipótesis explicativa y la verdad –o, al menos, aproximación a la verdad– de dicha hipótesis. Sin embargo, el único medio legítimo para *fundamentar* la convicción anterior parece ser la *apelación al privilegio*²⁰, esto es, la consideración según la cual los seres humanos –o, al menos, un subconjunto privilegiado de ellos, los científicos– están naturalmente predispuestos a dar con un conjunto de hipótesis en cuyo interior se encuentra la verdadera. Sólo esta especie de acto de fe con respecto a una facultad más *instintiva* que racional explicaría la alta probabilidad de que la hipótesis seleccionada como la mejor hipótesis explicativa de una evidencia dada sea verdadera, pero señala van Fraassen que resulta extremadamente difícil defender en la actualidad la idea de que exista tal capacidad para acertar con hipótesis verdaderas en vistas de los fracasos exhibidos por la historia de la ciencia real.

Volviendo a la tesis de unicidad levemente debilitada, creo que también resulta incompatible con la IME en la primaria versión de Harman: “uno infiere, a partir de la premisa de que una hipótesis dada proporcionaría una mejor explicación para la evidencia que cualquier otra hipótesis, la conclusión de que la hipótesis es verdadera” (Harman, 1965: 89). Ello deja abierta la posibilidad de que, en ciertos contextos, la premisa de que “una hipótesis dada proporcionaría una *mejor* explicación para la evidencia que cualquier otra hipótesis” no esté realmente disponible, ante lo cual resultaría perfectamente plausible suspender el juicio y no efectuar una IME en absoluto. Hasta aquí estaríamos satisfaciendo ‘Unicidad’ si entendemos que esta nos dice que *hay como máximo* una actitud racional que se puede asumir, lo cual supone que *podría no haber ninguna*. No obstante, Harman agrega que en general habrá muchas explicaciones que podrían explicar igualmente bien la evidencia y que, si bien en el caso ideal deberíamos ser capaces de rechazar todas las hipótesis alternativas antes de efectuar la inferencia que nos lleve a la ‘mejor

²⁰ Creo que la ‘tesis del privilegio’ fue brillantemente adelantada por Charles Peirce, quien describe la ‘abducción’ como una “capacidad para ‘adivinar’ correctamente”, un ‘misterioso poder adivinator’ que subyace a toda la investigación científica” (Peirce, 1958: 6.530). La apelación al privilegio es todavía más clara en la siguiente cita: “La sugerencia abductiva viene a nosotros como un destello. Es un acto de *insight*, aunque [se trata] de un *insight* extremadamente falible” (Peirce, 1958: 5.181).

explicación', en la práctica el gran inconveniente consiste en cómo definir 'mejor' de una manera objetiva. El autor adelanta el problema, pero no lo resuelve. De donde podemos suponer que valen aquí las mismas consideraciones vertidas sobre la IME liptoniana en relación con la existencia de ciertos contextos en los cuales la inferencia hacia varias buenas explicaciones compatibles con la evidencia estaría habilitada.

Resta analizar qué sucede con la 'Unicidad' así entendida en conexión con la subclase de la IME presentada por Bird: la IUE o inferencia holmesiana nos permitiría seleccionar *solo una* explicación potencial, dado que sería *la única* explicación potencial auténticamente consistente con la evidencia. Esto último parece presuponer no sólo la unicidad de White –como se adelantó–, sino también la de Feldman, puesto que –por definición– en el empleo de la IUE resultará como máximo una única explicación favorecida, dado que será la única que eludirá las refutaciones.

En resumen, sostengo que la mejor formulación de 'unicidad' es la de Feldman. Y esta última, conceptualmente, no está necesariamente presupuesta en la originaria IME de Gilbert Harman ni en la más reciente versión de Peter Lipton, aunque sí parece estarlo en la variante imaginada por Alexander Bird. De modo que la IUE de Bird requiere que valga unicidad (incluso en su versión débil), lo cual inhabilitaría *prima facie* que los DREP acontezcan. A continuación, analizaremos si la afirmación anterior se comprueba (o no) empíricamente con un caso tomado de la historia de la ciencia real.

Análisis de la IUE de Bird a la luz de los DREP: HIV y SIDA.

Un caso bastante reciente en la historia de la medicina que ilustraría la aplicación de la IUE de Bird es el descubrimiento del HIV y la causa del SIDA. La fase inicial involucró la identificación de un síndrome que requirió explicación. En junio de 1981 se publicó un reporte sobre la aparición de una forma rara de neumonía, la *pneumocystis*, en cinco hombres californianos homosexuales. La neumonía por *pneumocystis* solo se había observado en individuos que se habían sometido a terapias médicas que implicaban inmunosupresión. Al mes siguiente apareció un segundo reporte, discutiendo los casos de veintiséis hombres jóvenes homosexuales con el sarcoma de Kaposi, una forma inusual de cáncer de piel, normalmente encontrada solo en hombres mayores que están en sus setenta años aproximadamente y que son en general de origen mediterráneo. Además, cuatro de ellos tenían neumonía *pneumocystis* también. Poco después otros diez casos de *pneumocystis* fueron revelados en California. Como los centros para el control de la enfermedad (CDC=*Centers for Disease Control*) comentaron, "El aparente agrupamiento de la neumonía *pneumocystis carinii* y del sarcoma de Kaposi entre los hombres homosexuales sugiere un factor subyacente común" (Connor and Kingman, 1989: 14; citado por Bird, 2010: 349). La unión de síntomas en una manera indicativa de una causa común es un síndrome, en este caso inicialmente llamado GRIDS –*Gay-Related Immune Deficiency Syndrome*– y luego AIDS²¹ –*Acquired Immune Deficiency Syndrome*–.

¿Qué explica la existencia de este síndrome? ¿Qué causa el SIDA? Los investigadores consideraron las cuatro hipótesis que presentamos a continuación:

²¹ SIDA en lengua española.

H₁: Drogas recreativas. Inicialmente se sospechó de una tanda contaminada de ‘poppers’ (nitrato de amilo). Y luego se consideró que el uso excesivo de ciertas drogas recreativas, aun si no están contaminadas, podría deprimir el sistema inmune.

H₂: La muy alta incidencia de enfermedades de transmisión sexual familiar entre ciertos hombres sexualmente muy activos podría sobrecargar el sistema inmune y provocar su depresión. Esto también podría explicar la aparición del SIDA entre los usuarios de drogas intravenosas que compartieron agujas sucias.

H₃: Infección bacteriana –infección por medio de una bacteria, probablemente aun no descubierta–.

H₄: Infección viral –infección por medio de un virus, probablemente aun no descubierta–.

Cabe destacar que se supone la falsedad de las siguientes tesis adicionales:

T₁: No hay una causa común –*i.e.*, la unión de síntomas es enteramente accidental–.

T₂: No hay nada para explicar – la IME, en cualquiera de sus variantes, presupone que hay algo que requiere explicación. Si este no fuera el caso, no habría nada que cuente como mejor explicación–.

Bird llama “hipótesis nula” a T₂. El equivalente a la hipótesis nula en el ya célebre caso del Dr. Semmelweis²² –también analizado por Bird– sería la afirmación de que no hay una real diferencia médica en la cantidad de muertes por fiebre puerperal entre las divisiones I y II, sino que por mera casualidad las pacientes ingresadas en la división I eran individualmente más propensas a contraer la enfermedad. Semmelweis nunca pensó seriamente en esa posibilidad y el sentido común –como así también la opinión de los estadísticos modernos– nos confirma que en este caso la hipótesis nula no funciona. Del mismo modo,

el número de casos de síntomas raros, a menudo superpuestos, todos relacionados con un sistema inmune debilitado, y encontrados de muchos modos entre hombres homosexuales, significa que uno puede concluir en el caso del SIDA que la hipótesis nula es falsa. De modo que hay un síndrome genuino que necesita explicación (Bird, 2010: 349, la traducción es mía).

La pieza clave de evidencia que refutó las hipótesis vinculadas con los estilos de vida (H₁ y H₂) fue el descubrimiento del SIDA entre hemofílicos. En 1982, se encontró que varios hemofílicos padecían el síndrome, al igual que varias personas, tanto hombres como mujeres, que habían recibido transfusiones de sangre, incluido un bebé de veinte meses. Entre los donantes de la sangre que recibió el bebé había un hombre que desarrolló SIDA menos de un año después de la donación. Mientras que semejante evidencia apunta a una infección transmitida por sangre, también sirve para excluir las hipótesis H₁ y H₂, dado que ahora numerosos individuos estaban empezando a ser diagnosticados con SIDA, individuos que simplemente no consumían drogas ni tenían hábitos de sexo muy activo. En efecto, esta evidencia servía para refutar muy bien cualquier hipótesis relacionada con los estilos de vida, dado que no hay hábitos compartidos entre

²² Cf. Hempel, 1973: 16-20.

los hemofílicos, los hombres homosexuales y aquellos que reciben transfusiones de sangre, que no sean compartidos también por todos los demás. Por eso afirma Bird que “en mi opinión, es difícil pensar cualquier hipótesis compatible con la evidencia de los hemofílicos y los receptores de transfusiones que no tome a un agente infeccioso como la causa del SIDA” (*Ibid.*: 349).

Si en vez de tener dos hipótesis distintas (H_3 y H_4) tuviéramos una hipótesis más general, a saber, que “el SIDA es causado por un agente infeccioso”, entonces esa hipótesis es confirmada por medio de la refutación de la hipótesis nula y de todas las otras hipótesis inconsistentes con ella. Pero habiendo establecido que el SIDA es una enfermedad infecciosa, la siguiente tarea es identificar el tipo de infección. Los dos candidatos obvios son la infección bacteriana (H_3) y la infección viral (H_4). La evidencia ya obtenida descarta la infección bacteriana. Esto es así porque el producto sanguíneo utilizado por los hemofílicos, el agente de coagulación factor VIII, se obtiene de la sangre donada mediante un proceso que involucra, entre otras cosas, filtración. La filtración remueve las bacterias. De este modo, la hipótesis bacteriana puede ser excluida. Y con la hipótesis bacteriana refutada, es natural orientarse hacia la hipótesis viral. Sin embargo, uno podría preguntarse si algún otro agente infeccioso podría ser el responsable: no toda infección es bacteriana o viral (interpretando la disyunción en sentido excluyente); las otras posibilidades incluyen hongos, protozoos y parásitos multicelulares. Puede decirse que es concebible que algún tipo de agente filtrable aun no descubierto sea el responsable. Sabemos ahora que hay tales agentes.²³ Pero dado que la mayoría de ellos son similares a los virus (virus satelitales y viroides²³), podemos ignorar esta posibilidad.

Ahora bien, más allá de la refutación de todas las hipótesis explicativas que competían con H_4 , ¿qué fue lo que estableció definitivamente que es un virus el que causa el SIDA? La respuesta se halla en el aislamiento de un virus particular, el LAV (*lymphadenopathy AIDS-associated virus*), por parte de Luc Montagnier en 1983, rebautizado como HIV tres años más tarde. A su debido tiempo, se mostró que el HIV satisface los postulados de Koch²⁴ con respecto al SIDA.

Según Bird, lo que el caso del SIDA muestra es que la identificación de la hipótesis explicativa correcta acaece en virtud de la refutación de las principales rivales. Y agrega que, aunque la metodología pueda tener un sabor popperiano, la epistemología no lo tiene, dado que el proceso de eliminación aumenta la probabilidad de la hipótesis que permanece irrefutada²⁵. La evidencia epidemiológica descarta hipótesis tales como la sobrecarga del sistema inmune por drogas; en efecto, resulta refutada cualquier hipótesis que no permita la infección transmitida por sangre. Todo ello aumenta la probabilidad de que la causa del SIDA sea una bacteria o un virus. El hecho de que el agente infeccioso sea filtrable descarta la infección bacteriana, incrementando la probabilidad de que la causa sea un virus. Sin embargo, esto no establece la hipótesis viral con certeza, dado que la evidencia es consistente con agentes subvirales como causa. Pero el hecho de que tales agentes sean raros significa que la hipótesis del virus tiene una alta probabilidad, lo que alentó a Montagnier a buscar el virus directamente.

²³ Se sabe que típicamente requieren la presencia de un virus real para replicarse.

²⁴ Los postulados de Koch son principios usados para establecer que un cierto agente infeccioso es la causa de una enfermedad dada.

²⁵ Recordemos que el rechazo popperiano de cualquier forma de inducción le impide sostener que las hipótesis corroboradas sean probables, dado que, de acuerdo con la ley del cálculo de probabilidades, la probabilidad de una hipótesis de carácter universal siempre tiende a 0 (cero).

¿Qué hay de los DREP en casos como el analizado? Es claro que no pueden tener lugar dichos desacuerdos, al menos no al final del proceso que da como resultado que sólo una explicación resulta auténticamente consistente con la evidencia, dado que todas las demás van siendo progresivamente descartadas. Resta averiguar si habrá más casos que avalen la aplicación legítima de una IUE, caso límite de la IME. Pues Bird sólo nos ofrece dos casos –el de la fiebre puerperal en el hospital general de Viena y el del descubrimiento del HIV– tomados de la historia de la Medicina. Esta última, se podría objetar, no es una ciencia *stricto sensu*, sino más bien una *técnica científica* que se vale de las contribuciones de varias disciplinas científicas: la química, la biología, la física, etcétera.

> Conclusión

El interrogante que titula este trabajo podría responderse negativamente sólo si se tuviera en cuenta que existe un caso límite de la IME –la IUE de Bird– que inhabilita desacuerdos racionales entre pares epistémicos o, dicho de otro modo, que requiere que valga la tesis de unicidad (en cualquiera de las versiones analizadas, inclusive en la débil de Feldman). No obstante, sabemos que la IUE es un caso más bien raro de IME que no abunda en la práctica científica corriente. En ella son mucho más usuales las aplicaciones de inferencias a la mejor explicación que admiten que pueda haber desacuerdos racionales entre pares epistémicos, lo cual es perfectamente consistente tanto con la IME inicialmente presentada por Harman como con la IME liptoniana comúnmente tomada como modelo en la actualidad.

La IME de Lipton concede que, en ciertas circunstancias, ninguna explicación merezca ser favorecida –por no ser ‘lo suficientemente buena’– e incluso que en otros contextos más de una explicación pueda ser utilizada (como en el caso de la explicación mecanicista y de la explicación funcional ya analizado). De donde se sigue que la IME de Lipton no presupone la unicidad ni siquiera en su versión débil y, por ende, habilita los desacuerdos racionales entre pares epistémicos.

Por otra parte, la elección de la mejor explicación en el modelo liptoniano no es tan sencilla si se recuerda que hay diferentes criterios que pueden tomarse en consideración (y sobre los cuales no tiene por qué haber consenso en la comunidad científica). El criterio más popular es el de la probabilidad o verosimilitud (*likeliness*). Es decir, se trataría de elegir la explicación mejor respaldada por la evidencia, lo cual la volvería más probablemente verdadera. Sin embargo, Lipton propone evaluar otra cualidad de las explicaciones, que podría ser un nuevo criterio de selección sumamente relevante a sus ojos: la *loveliness*, su encanto explicativo. Esta noción apuntaría ya no a la verdad, sino a la comprensión que una explicación nos podría brindar (si fuera, además, verdadera), la *comprensión potencial*. A menos que se elucide la noción de “*loveliness*” en términos completamente objetivos –cosa que hasta el momento no se ha logrado–, deberemos acordar en el hecho de que la explicación que a usted le brinde la mayor comprensión potencial no tiene por qué coincidir con la mía. De modo que vislumbramos aquí una razón adicional para que los desacuerdos entre pares acontezcan en las aplicaciones concretas de la IME.

> Bibliografía

- » Ayer, A. J. (1936). *Language, Truth and Logic*, London: Penguin.
- » Azar, R. M. (2017). “¿Es correcta la identificación entre ‘Abducción’ e ‘Inferencia a la Mejor Explicación’?”, *Contrastes*. Revista Internacional de Filosofía, vol. XXII, Nº1, pp. 7-17.
- » Bergmann, Michael (2009). “Rational Disagreement after Full Disclosure”. *Episteme* Vol. 6, Nº3, pp. 336-353.
- » Bird, A. (2010). “The Epistemology of Science – A Bird’s-Eye View”, *Synthese*, Nº 175, pp. 5–16.
- » Bird, A. (2007). “Inference to the Only Explanation”, *Philosophy and Phenomenological Research*, Vol. 74, Nº 2, pp. 424-432.
- » Carnap, R. (1950). *Logical Foundations of Probability*, Chicago: Chicago University Press.
- » Christensen, D. (2007). “Epistemology of Disagreement: The Good News”, *Philosophical Review*, Nº 116, pp. 187-217.
- » Conan Doyle, A. (1953). “The sign of four”. En: *The complete Sherlock Holmes*, (Vol. I). New York: Doubleday.
- » Douven, I. (2009). “Uniqueness revisited”. *American Philosophical Quarterly*, Vol. 46, Nº4, pp. 347 - 361.
- » Feldman, R. (2007). “Reasonable Religious Disagreements,” En L. Antony (ed.) *Philosophers Without Gods: Meditations on Atheism and the Secular Life*, Oxford: Oxford University Press, pp. 194–214
- » Feldman, R. (2006). “Epistemological Puzzles about Disagreement,” En S. Hetherington (ed.) *Epistemology Futures*, Oxford: Oxford University Press, pp. 216–236.
- » Harman, G. (1986). *Change in View: Principles of Reasoning*. Cambridge: The MIT Press.
- » Harman, G. (1965). “The Inference to the Best Explanation,” *Philosophical Review*, Vol. 74, Nº1, pp. 88–95.
- » Hempel, C. (1973). *Filosofía de la ciencia natural*. Madrid, Alianza editorial.
- » Hempel, C. (1965). *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science*, New York: Free Press.
- » Hempel, C. 1945). “Studies in the Logic of Confirmation,” *Mind* 54:1–26, pp. 97–121.
- » Huemer, M. (2011). “Epistemological egoism and agent-centered norms”. En Trent Dougherty (ed.), *Evidentialism and its Discontents*. Oxford University Press., pp. 17-40.
- » Kuhn, T. (1977). “Objectivity, Value Judgment, and Theory Choice”, In his *The Essential Tension*, Chicago: University of Chicago Press, pp. 320–339
- » Lipton, Peter (1991). *Inference to the Best Explanation*, London, Routledge.
- » Persson, J. (2007), “IBE and EBI. Explanation before inference”, en: Persson, Johannes e Ylikosky, Petri (2007), *Rethinking Explanation*, The Netherlands: Springer, pp. 137-148
- » White, R. (2005). “Epistemic Permissiveness,” *Philosophical Perspectives*, Vol. 19, Nº1, pp. 445–459.