

*Explicación
teórica*

**Explicación contextualista
y ratificación etnocéntrica
de los criterios
de justificación***

Universidad de Buenos Aires.

EN EL PRESENTE TRABAJO voy a defender la concepción conversacionalista de la justificación en la caracterización que de la misma ha hecho y sigue haciendo Richard Rorty. La defensa en cuestión será literal, quiero decir, me ocuparé de cuestionar algunas objeciones que se le han formulado. En nuestro coloquio del año 1999 realizado en Puebla, dediqué buena parte de la exposición a enfrentar argumentos que pretendían cuestionar al conversacionalismo desde posiciones, llamémoslas con rapidez, “realistas”¹. En esta ocasión me detendré en un grupo de objeciones provenientes del campo pragmatista. Vertebraré la primera parte de la exposición tomando como referencia la crítica a Rorty sostenida por Susan Haack, para ir luego incorporando las observaciones que Hilary Putnam, Akeel Bilgrami y Jürgen Habermas le han presentado al proyecto por defender.

Una manera de presentar el modo en que desde la epistemología tradicional se concibe la tarea de delinear una teoría de la justificación es como una empresa apriorística que debe ofrecer respuestas a las siguientes dos preguntas: “¿Qué es lo que debe ser considerado como otorgando apoyo racional a una creencia?” y “¿Qué relación existe entre el hecho de que una creencia tenga tal apoyo (es decir, que esté justificada) y la probabilidad de que la misma sea verdadera?”. Susan Haack (1993, cap. 1) ha llamado *proyecto de explicación* al intento de responder a la primer pregunta y *proyecto de ratificación* a la búsqueda de respuestas a la segunda.

* Este trabajo fue escrito durante el período en que el autor fue beneficiado con una beca para terminar su doctorado en Argentina otorgada por la Fundación Antorchas. Quiero agradecer a la Fundación por dicho apoyo.

1. Véase Penelas (2001a).

Uno de los más firmes intentos de rechazo a dicho proyecto teórico es el que viene desplegando Richard Rorty desde la publicación de *La filosofía y el espejo de la naturaleza*. Buena parte de su prédica a favor del abandono de la epistemología, tal como fue concebida en la modernidad, se presenta en realidad como una crítica a los presupuestos fundacionalistas que condicionan a dicha empresa. Ha sido Susan Haack (*ibid.*, cap. 9), una vez más, la que ha señalado que hay que ver a la impugnación que hace Rorty de la epistemología como oponiéndose a la vez a tres tipos de fundacionalismo, sin distinguir los matices que hay entre ellos. El primer tipo de fundacionalismo es una forma de llevar a cabo el proyecto de explicación, el segundo es la tesis que origina la relevancia del proyecto de ratificación, y el tercero consiste en una caracterización metafilosófica de ambos proyectos. Llamaré, con ligereza, a estas tres formas de fundacionalismo *experiencialismo*, *veroteleologismo* y *apriorismo*, respectivamente. La caracterización que hace Haack de las tres tesis que Rorty confunde bajo el rótulo de *fundacionalismo* es la siguiente:

experiencialismo: hay dos tipos de creencias, las creencias básicas que se justifican por la experiencia y las creencias derivadas que se justifican sobre la base de las creencias básicas.

veroteleologismo: los criterios de justificación no son convencionales sino que tienen una base objetiva, siendo satisfactorios sólo si son indicativos de *verdad*.

apriorismo: la epistemología es una disciplina *a priori* en tanto la explicación de los criterios de justificación es una iniciativa analítica, y la ratificación de los mismos depende de una prueba *a priori* de su carácter indicativo de *verdad*.

Lo primero que hay que observar es en qué sentido estas tesis son independientes. Según Haack, la tesis veroteleologista no implica ni la tesis apriorista ni la experiencialista, y la tesis apriorista no implica a la tesis experiencialista. Su propia posición consiste en negarle carácter apriorístico a la epistemología (contra el *apriorismo*), rechazar el *experiencialismo*, adoptando una posición que se aleja también del *coherentismo* (a la que llama *fundherentismo*), y abrazar la tesis veroteleologista como inescindible de una caracterización adecuada de la idea de *justificación*. No voy a detenerme en las razones que da la autora para mostrar qué relaciones inferenciales se dan entre las tres tesis, ni en la caracterización pormenorizada de su propia posición. Sólo prestaré atención a su descripción de los alcances y límites del rechazo rortiano a las tres formas de fundacionalismo.

Haack concuerda con Rorty en su rechazo al experiencialismo y al apriorismo. Señala que el sustento argumentativo para rechazar al primero lo encuentra Rorty en la obra de Sellars, y para abandonar el segundo, en la obra de Quine.

Todo lo que hace Rorty es apoyarse en la objeción sellarsiana según la cual el experiencialismo confunde las razones con las causas, y en la crítica quineana a los dos dogmas del empirismo, la cual conduce a alguna forma de naturalización de la epistemología (que en Rorty se presenta de una manera ampliada, incorporando a las ciencias sociales dentro del *corpus* de disciplinas que deben ser atendidas a la hora de explicar el fenómeno cognitivo). El abandono de Haack de las primeras dos formas de fundacionalismo se da por razones distintas a las que encuentra en Rorty. Pero tampoco me detendré en esas diferencias, sino en el punto fuerte de discrepancia entre ambos, que se despliega en dos aspectos: en qué tipo de alternativa ofrecer al experiencialismo, y en la adopción por parte de Haack y el abandono por parte de Rorty del veroteleologismo. La idea de Haack se resume en que las razones para abandonar el apriorismo y el experiencialismo no bastan para declarar la defunción de la epistemología, pues el proyecto explicativo se mantiene en pie en tanto haya alternativas viables al experiencialismo (el coherentismo, que ella también rechaza, y el fundherentismo que defiende), y el proyecto de ratificación no es más que lo que todo veroteleologista debe llevar a cabo, siendo el veroteleologismo ineludible.

La posición cuestionada por Haack es el *conversacionalismo* sostenido por Rorty. El conversacionalismo debe ser visto, según la autora, como la suma de dos tesis de distinto nivel: una tesis contextualista sobre la explicación de los criterios de justificación y una tesis convencionalista sobre la ratificación de tales criterios. Las tesis son enunciadas de la siguiente manera:

contextualismo: “A está justificado a creer que p si con respecto a p A satisface las pautas epistémicas de la comunidad epistémica a la que pertenece” (se trata pues de una caracterización de los criterios de justificación que contrasta con el experiencialismo, y que a la vez es incompatible con el apriorismo, pues la explicación de los criterios sólo podrá ser una explicitación de los criterios implícitos en las prácticas de justificación y, por ende, una tarea que no tiene el rasgo de lo apriorístico);

convencionalismo: los criterios de justificación son convencionales, no tiene sentido preguntar cuáles son los criterios correctos de justificación, cuáles son realmente indicativos de la probable verdad de una creencia (se trata pues de una tesis metaepistemológica que contrasta con el veroteleologismo, y desde la cual el proyecto de ratificación adquiere el rasgo de la ininteligibilidad).

En la manera como Haack cuestiona al conversacionalismo rortiano se pueden distinguir tres pasos. Uno en contra del contextualismo, otro en contra del convencionalismo, y otro en contra del conversacionalismo en su conjunto. Su

objección al contextualismo en realidad no es directa; lo único que dice es que Rorty presenta su posición contextualista como oposición al experiencialismo, una vez aceptada la crítica de origen sellarsiano, sin advertir que hay alternativas al mismo que Rorty no contempla y que no se ven afectadas por el argumento de Sellars, esto es, el coherentismo y el fundherentismo de la propia Haack. Como puede verse, el argumento de Haack hasta aquí no alcanza a cuestionar al contextualismo; sólo recuerda que el mismo no se sigue del rechazo al experiencialismo. Lo siguiente es señalar que el contextualismo sin el convencionalismo es una tesis trivial, y que, por tanto, si bien no implica, sí motiva fuertemente la adopción de la posición convencionalista, y que, consiguientemente, un argumento en contra del convencionalismo supone una buena razón en contra también del contextualismo.

Ahora bien, ¿cuál es la crítica que hace Haack a la adopción por parte de Rorty del convencionalismo? Su idea es que Rorty no presenta más razones para rechazar el veroteleologismo que su crítica a la concepción correspondentista de la verdad, sin dar cuenta de que entre el más crudo correspondentismo y su particular deflacionismo hay muchas otras teorías de la verdad frente a las que no ofrece argumento alguno. Con lo cual, el abandono del veroteleologismo está a la espera de argumentos que descarten a cada una de las teorías pertinentes. Esta observación de Haack es falsa en dos sentidos. En primer lugar, porque buena parte de las teorías de la verdad enunciadas por ella² son explícitamente rechazadas por Rorty (o asimiladas a su posición) sobre la base de argumentos propios o, paradigmáticamente, de Davidson³. Y, lo que es más importante, la afirmación de Haack es falsa, porque Rorty tiene un argumento directo contra el veroteleologismo a través del cual se cuestiona fuertemente lo que está a la base de dicha posición, esto es, la concepción de la verdad como una meta de la investigación o de la justificación⁴. El argumento puede presentarse como sigue:

Premisa 1: Una meta no reconocible al ser alcanzada no puede ser intencionada como meta.

Premisa 2: En el contexto de la investigación podemos reconocer cuándo tenemos por justificada una creencia, pero no cuándo una creencia es verdadera.

-
2. Haack identifica los siguientes grupos de teorías de la verdad como los no contraargumentados por Rorty: teorías pragmatistas (Peirce), mínimamente realistas (Ramsey, Tarski), fuertemente realistas (Wittgenstein I, Russell, Austin).
 3. Sería fatigoso señalar todas las citas pertinentes que mostrarían estos rechazos explícitos. Como muestra puede verse Rorty (1991), en donde varias de las opciones señaladas por Haack son cuestionadas.
 4. El argumento puede rastrearse, entre otros textos, en Rorty (1998).

Conclusión 1: La verdad no es una meta de la investigación; sólo la justificación puede serlo.

Conclusión 2: En el contexto de la investigación, la verdad no difiere prácticamente de la justificación.

Premisa 3: Si algo no hace una diferencia en la práctica, entonces no debería hacerla tampoco en la filosofía.

Conclusión 3: No hay diferencia filosófica entre verdad y justificación⁵.

El argumento de Rorty (acerca del cual volveré más adelante en relación con algunas objeciones que se le han presentado), da lugar al tipo especial de deflacionismo con respecto a la verdad defendido por el autor, esto es, la posición que niega que haya usos del término verdadero que no sean parafraseables en términos de justificación. Rorty aborda su discusión acerca del concepto de *verdad* en términos de los usos atribuibles o no atribuibles a dicha palabra. Su postura deflacionista conjuga la tesis de que no hay ninguna propiedad aplicable a una oración dada que sirva para *explicar* que dicha oración es verdadera con la idea de que entre los usos reconocibles de *verdad* no debemos incluir un uso *explicativo*. Así, *verdad* no necesita de una definición que nos brinde una propiedad común subyacente a todas las oraciones verdaderas que permita explicar su verdad; de lo que se trata es de especificar los usos que el concepto de verdad posee, agregándose a continuación la tesis de que *verdad* no tiene usos explicativos. Rorty señala tres usos asociados al concepto de verdad:

“un uso de respaldo (*endorsing use*)

un uso precautorio (*cautionary use*), en observaciones tales como “Tu creencia de que S está perfectamente justificada, pero quizás no es verdadera”, que nos recuerdan que la justificación es relativa a, y no es mejor que, las creencias citadas como base para S, y que dicha justificación no es garantía de que las cosas vayan bien si tomamos a S como “regla para la acción” (definición de *creencia* de Peirce);

un uso desentrecomillador: para decir en el metalenguaje cosas de la forma: “S es verdadera si y sólo si...” (cf. Rorty, 1991, p. 128).

Tales usos son concebidos por Rorty, a su vez, como claramente parafrásticos. El uso desentrecomillador es claramente una explicitación de la capacidad del predicado de verdad como recurso para la formulación de parafrasis o abreviaturas. Por su parte, los usos laudatorio y precautorio son reformulables en términos de

5. Debo la inspiración de esta presentación del argumento de Rorty a Daniel Kalpokas.

mera justificación. El uso laudatorio no es más que una manera enfática de decir que una oración está justificada. El uso precautorio no es más que una manera de decir que lo que está justificado en el presente puede no estarlo en el futuro. Como dice Rorty,

el uso precautorio es usado para contrastar audiencias poco informadas con audiencias mejor informadas, audiencias pasadas con audiencias futuras [...] el punto de contrastar la verdad y la justificación es simplemente recordarse a uno mismo que puede haber objeciones que no se le han ocurrido todavía a nadie (Rorty, 2000, p. 4).

Hasta aquí entonces hemos visto cómo las objeciones al contextualismo y el convencionalismo presentadas Haack no alcanzan a dar en el blanco pues sólo se basan en pedir más argumentos, siendo que en lo sustancial Rorty no presenta los huecos argumentativos que la autora pretende. Prestemos atención entonces a la objeción directa que presenta Haack al conversacionalismo. Lo que se cuestiona es que el conversacionalismo redunde en caracterizar un sujeto epistémico relativista y cínico. Relativista, pues todas las pautas comunitarias de justificación serán juzgadas como estando en un pie de igualdad desde un punto de vista evaluativo, y cínico porque aceptará siempre las propias pautas de justificación sin aceptarlas realmente. Haack remata diciendo que uno no puede involucrarse coherente y completamente, sin cinismo, en la práctica de justificar creencias, si uno concibe a la misma como enteramente convencional. Porque creer que *p* es aceptar *p* como verdadera.

Lo primero que debería decirse frente a esta acusación de relativismo es lo que el mismo Rorty le advierte a Putnam cuando este último le hace la misma acusación: el relativismo se autorrefuta, pero la posición adoptada no es la de que todo canon es igualmente bueno, sino la de que son nuestros cánones de justificación los cánones adecuados. El conversacionalismo no debe pues presentarse como lo hace Haack, como la conjunción de la tesis contextualista con la tesis convencionalista, sino como la conjunción del contextualismo con el etnocentrismo, el cual sí se presenta como una tesis sobre la ratificación (aunque ya no en términos de *verdad*) que puede enunciarse así: los criterios correctos de justificación son los *nuestros*, donde la corrección tiene que ver con que dichos criterios conducen a consensos amplios y estables. En efecto, para Rorty la noción de *justificación* está estrechamente relacionada con la de *consenso*, aunque no en el sentido de que deba analizarse una en términos de la otra, sino en el sentido de que es ininteligible la idea según la cual puede decirse de una oración que no está justificada aun cuando la totalidad de los hablantes de una comunidad la aceptan. La justificación es independiente del consenso mayori-

tario pero no del consenso global. ¿Cómo podría serlo? Si justificar no es otra cosa que involucrarse en la práctica de dar y pedir razones, ¿cómo puede decirse que una oración no está justificada cuando tal juego se ha cerrado en virtud de que ya nadie disputa sobre dicha oración, quedando quienes la defienden libres de la obligación de ofrecer respaldos explícitos? Si justificar es justificar frente a los miembros de mi comunidad, ¿cómo puede tener sentido la idea de que una oración que todos aceptan en determinadas ocasiones no está justificada cuando tal consenso se da? Ahora bien, el que la identificación de una comunidad en la que son aceptadas globalmente ciertas creencias nos imponga calificar ese contexto de justificación como uno donde tales creencias están justificadas para los miembros de esa comunidad, no nos impide, en virtud de que nosotros no aceptamos tales creencias, evaluar los estándares de justificación de la comunidad ajena como peores que los nuestros.

Haack no advierte esta apuesta etnocentrista de Rorty, y se pierde entonces de apuntar al aspecto más controvertido del conversacionalismo rortiano. Putnam, por su parte, sí lo ha visto y ha ofrecido una crítica que todo defensor de la propuesta rortiana no puede soslayar, esto es, la idea de que contextualismo y etnocentrismo son incompatibles, pues su conjunción conduce a una noción ininteligible de reforma epistémica⁶. Rorty no puede sino concebir la idea de que hay mejores cánones de justificación que otros en términos de “lo que llega a parecernos mejor a nosotros”, donde “nosotros” debe pensarse como “nosotros en la mejor versión que podamos concebir de nosotros mismos”. Ahora bien, para Putnam, dicha posición rortiana con respecto a cómo evaluamos estándares de justificación alternativos hace ininteligibles los proyectos reformadores en una comunidad etnocéntrica, pues es interno a nuestro concepto de *reforma* el que lo bueno y lo malo sean independientes de lo que nos parece bueno o malo, pues siempre nuestro cambio de creencias nos va a parecer bueno. La posición de Rorty sólo permite la contraposición de distintos estándares de justificación, pero no la evaluación de estándares ajenos al nuestro.

El punto de Putnam es que, para Rorty, el cambio conceptual no es racional y que, entonces, desde el etnocentrismo no puede hablarse de reforma. Lo que Rorty responde es que obviamente no puede hablarse de reforma *sub specie aeternitatis*, pero podemos desde nuestra comunidad hacer juicios evaluativos. Podemos reconocer distintas comunidades de justificadores, y evaluarlas como peores que la nuestra.

6. He presentado esta crítica de Putnam, con respuestas a la misma, diferentes y análogas a las que aquí desarrollo, en Penelas (2001b) y Penelas (2003).

El punto de Putnam es que, si justificación no es más que consenso, entonces toda comunidad del pasado en la que había consenso acerca de p debe ser vista como teniendo por justificada p , pero entonces no puede decirse que nuestra comunidad que acepta consensualmente no p sea mejor que la otra, pues en ambas habría justificación. ¿En qué sentido sería mejor?

Frente al punto sólo voy a decir tres cosas. En primer lugar, que según Putnam (y, si no me equivoco, según Haack también) los estándares de justificación dependen de los intereses y valores desplegados históricamente por las comunidades humanas, lo cual no veo cómo hacer compatible con lo que se desprende de su objeción a Rorty (que p no estaba justificada para la comunidad del pasado en la que era aceptada globalmente). Además, me parece que, si algo nos ha dejado Kuhn, es, al menos, cierta intuición a favor de la idea de que, aunque falsa, p estaba justificada. Lo segundo es que deberíamos precisar en qué sentido hablamos de “mejor” cuando, con Rorty, decimos que podemos decir que p estaba justificada en el pasado (en virtud del consenso) y que, sin embargo, nuestro modo de justificación es mejor. Apelo aquí, y como mera tentativa a desarrollar, a una idea de Peirce. Cuando el supuesto padre del pragmatismo señala los cuatro modos de fijación de creencia (modos de justificación), la evaluación entre ellos es pragmática, se da en términos de estabilidad (*cf.* Peirce, 1931-1958, Vol. V §§ 358-387). La evaluación positiva que hace Peirce del método científico frente a los métodos de la tenacidad, de la autoridad y del apriorismo no se da en términos de concordancia con condiciones epistémicas ideales. El método científico es el mejor porque es el más estable. Los otros tres métodos son inferiores porque tarde o temprano conducen a la inestabilidad, la cual es más difícil de darse si se sigue lo que Peirce llama método científico. La explicación de esa diferencia parece ser pensada en Peirce en clave sociológica o psicológica⁷. Con ese texto de Peirce como trasfondo, quizás podríamos defender el punto de vista etnocéntrico, apelando a esta idea pragmática de estabilidad y decir que, aunque justificadas, las creencias de nuestros antepasados se basaban en estándares de justificación peores que los nuestros, pues los nuestros son más estables, en el sentido de que generan consensos más estables. Quiero sólo señalar esta apelación a la idea de estabilidad como un camino posible por desarrollar y precisar a fin de proporcionar desde el etnocentrismo una defensa de la noción de evaluación interestándares más vigorosa que la del propio Rorty. Pero no puedo ofrecer más detalles al respecto, por lo menos por ahora.

7. Una sugerencia a favor de leer a Peirce a través de dicha clave puede encontrarse en el mismo Rorty (1995, p. 152).

En tercer lugar, quiero echar mano a la idea de Brandom según la cual la justificación se conforma a una estructura *default*-desafío (que no es sino una versión de la idea peirceana de que, frente al cartesianismo, debe pensarse la empresa cognitiva bajo el modelo creencia-duda y no bajo el modelo duda-creencia). Si atendemos a esta idea, en conjunción con la afirmación rortiana de que la empresa justificadora no es para nosotros, en cuanto animales sociales, una opción (así como no lo es la respiración), podemos terminar de responder a la objeción de Putnam y a la vez a la acusación de cinismo por parte de Haack. La idea de Brandom (1994, cap. 4) es que uno puede estar justificado aun si no ha atravesado un proceso efectivo y explícito de justificación, siendo éste el caso en un amplio margen de situaciones, en las cuales nuestras creencias, al no ser desafiadas, están claramente justificadas *by default* (recordando una vez más que *desafío* debe ser entendido peirceanamente, esto es, como desafío real y no fingido, como en la duda cartesiana). Ésta es una idea que Rorty abraza con gusto, y acuña la noción de *léxico último* para referirse justamente al conjunto de términos relevantes en el *corpus* de creencias que una comunidad adopta sin necesidad de justificación explícita en virtud de que no son desafiadas. Es razonable pensar que el léxico último de una comunidad está conformado en parte por las creencias mismas sobre adecuación racional, y que no serán desafiados y, por tanto, serán correctos etnocéntricamente en la medida en que dentro de nuestra comunidad dichos criterios de justificación nos conduzcan a consensos en la práctica de dar y pedir razones explícitamente, esto es, nos conduzcan a nuevas creencias no desafiadas. El que podamos observar comunidades donde otros cánones de justificación están vigentes y no son desafiados no nos impide juzgarlos como inadecuados, siempre y cuando no lleguen a desafiar nuestro cuerpo de creencias acerca de la justificación. Identificar otra comunidad de justificación es compatible con juzgarla como peor que la nuestra, esto es, como produciendo consenso en torno a oraciones que nosotros no aceptamos y que, por tanto, no consideramos verdaderas. Esto es así sólo si nuestro cuerpo de creencias básico sobre qué es la justificación no ha sido desafiado. Para un etnocentrista, la existencia de un consenso ajeno no alcanza para desafiar nuestras creencias. Es por eso que no hay incompatibilidad entre advertir que hay otros cánones de justificación que generan consensos y mantener los propios, pues la mera advertencia no basta para desafiarnos realmente. Es por eso que la acusación de cinismo no alcanza al conversacionalista, pues no se da el caso de que el investigador etnocentrista afirma lo que ya ha negado; el reconocer otros estándares de justificación no implica *ipso facto* el cuestionamiento de los propios. Justamente, en tanto etnocentrista, el investigador en cuestión no puede ser cínico; juzgará como falsas buena parte de las creencias consensuadas globalmente en comunidades extrañas, aun cuando en tanto contextualista dirá que las mismas están justificadas en esa comunidad.

Viene a cuento en este punto apelar a la distinción hecha por Robert Fogelin y retomada por Michael Williams entre dos planos de la justificación que deben estar involucrados en la atribución de conocimiento: el de la responsabilidad epistémica y el de la fundamentación adecuada. Williams caracteriza dichos planos en los siguientes términos (*cf.* Williams, 1999, p. 187):

responsabilidad epistémica: una persona está justificada en creer una proposición si ha formado o sostiene dicha creencia de una manera responsable.

fundamentación adecuada: la creencia de una persona está justificada si hay fundamentos que hacen probable que la proposición creída sea verdadera.

Rorty podría adherir a la idea de que para atribuir conocimiento debemos contemplar ambos planos, siempre y cuando consideremos la idea de fundamentación adecuada en términos etnocéntricos, esto es, remplazando “hay fundamentos que hacen probable que la proposición creída sea verdadera” por “sería aceptada globalmente en nuestra comunidad en virtud de nuestros estándares de justificación efectivizados sin distorsiones, esto es, sería aceptada por nosotros en nuestra mejor versión”. No parece ahora tan conflictiva la idea de considerar que los miembros de cierta comunidad están justificados al afirmar las creencias que aceptan consensualmente y que nosotros negamos. Lo que hacemos es atribuirles responsabilidad epistémica sin atribuir fundamentación adecuada a sus creencias.

Alguien podría decir que en realidad el problema de un etnocentrista claramente no es el cinismo, es decir, el fingir adoptar lo que no adopta, sino el dogmatismo, es decir, su incapacidad de dejar que sus adopciones sean desafiadas. Habría que señalar aquí que, como bien ha anotado Barry Allen, el etnocentrismo es, en Rorty, sostenido por el liberalismo. Dice Allen que

nuestro etnocentrismo (el de una comunidad democrática) es diferente al de todos los demás. Cuando *nosotros* somos etnocéntricos, no somos etnocéntricos. Cuando somos fieles a nuestras tradiciones, estamos atentos a otras tradiciones; cuando estamos interesados en nosotros mismos, estamos interesados en lo que es nuevo y diferente, felices de acomodarnos (lo más que nos sea posible) y aprender de ello (Allen, 2000, p. 224).

Se ve así la indisolubilidad que tiene en Rorty su posición epistemológica de su posición política, presentándose así como un fiel exponente de la tradición antirrepresentacionalista liberal inaugurada por Dewey. Frente al peligro dogmático que entraña el etnocentrismo debería decirse que no hay tal peligro porque queda disuelto por el *ethnos* liberal (aunque la apertura a otras comunidades que tal *ethnos* nos impone no debe llevar a pensar, como diremos más adelante, que

nos vemos a nosotros mismos como obligados a ser interlocutores universales). A la afirmación de que el etnocentrismo se vuelve inocuo en una cultural liberal yo creo que Rorty podría también preguntar retóricamente, “¿alguien más que un liberal podría ser genuinamente contextualista y etnocentrista?”. Clarificar por qué ésta sería una pregunta retórica para Rorty me llevaría a otra exposición completa, que les voy a quedar debiendo.

Hasta aquí, entonces, en lo que respecta a las objeciones al conversacionalismo que señalan el relativismo y el cinismo de la posición. En estas últimas páginas quiero responder a algunas objeciones que desde el campo pragmatista se le han hecho al deflacionismo de Rorty con respecto a la verdad, según las cuales la verdad es la meta de la investigación, y por tanto no es un mero recurso expresivo de paráfrasis de juicios que hacen predicaciones de justificación.

Desde el pragmatismo defendido por Bilgrami⁸, la epistemología debe asumir el punto de vista de la primera persona, y al hacerlo no puede dejar de advertir que la verdad es una norma y una meta de la investigación. Bilgrami no hace sino extraer esta conclusión del modelo creencia-duda de Peirce. Desde esta posición se asume que hay un conjunto de creencias que constituyen nuestra concepción del mundo y que nos sirven de estándares para la determinación de la verdad de las demás creencias. Tal grupo de creencias no necesita justificación en la medida en que las mismas son sostenidas sin dudas. Así la idea de justificar las creencias acerca de las cuales no tenemos dudas y que constituyen nuestra concepción del mundo carece de sentido desde el punto de vista epistemológico. La idea de justificación se pone en juego entonces sólo cuando se conmueve nuestra concepción del mundo y se da el proceso de cambio en la misma. Así, para Bilgrami, la verdad se aplica a las creencias que están fuera de duda, a la justificación a los cambios en nuestras creencias que sostenemos sin dudas. De modo que la meta de la investigación no es la justificación, sino la verdad, pues lo que el investigador quiere es saber cuál de las creencias acerca de las que tiene dudas puede incorporarse a su concepción del mundo, es decir, cuál es verdadera. Lo que pretende Bilgrami es cuestionar la premisa 2 del argumento de Rorty, señalando que en el contexto de la investigación (desde el punto de vista de la primera persona, que es el punto de vista relevante) sí podemos reconocer cuáles son las creencias verdaderas, pues éstas son las que constituyen nuestra concepción del mundo, y para ellas no necesitamos justificación alguna pues no dudamos de ellas.

El argumento de Bilgrami me parece sencillamente sorprendente. En primer lugar, como ya he dicho, Rorty (y más explícitamente Brandom) también asume que la

8. Para la crítica de Bilgrami a Rorty, véase Bilgrami (2000).

investigación acaece dentro de un marco de creencias que no son puestas en duda. Y, en virtud del uso laudatorio de *verdad*, no tendría problemas en decir que las tomamos como verdaderas, pues, junto con Brandom, considera a estas creencias como justificadas *by default*, justificadas en la práctica, y una marca de dicha justificación tácita se da en el consenso global no desafiado que soportan las mismas. Así, hasta aquí la diferencia con Bilgrami no parecería ser más que terminológica. Pero Bilgrami parece querer decir más, quiere decir que la verdad es la meta de la investigación. Ahora bien, en su planteamiento, si la verdad queda identificada con lo no cuestionado, entonces el objetivo de la investigación no es la verdad sino el acuerdo. Éste es el objetivo posible; la indudabilidad es un resultado del no desafío por parte de nadie. Bilgrami tiene razón en que el uso cautelar de *verdad* no tiene efectos perlocucionarios en el investigador, pues el señalamiento de que cualquiera de las creencias que tenemos por verdaderas puede ser falsa es una observación general incapaz de desafiar nuestra concepción del mundo (es justamente la crítica de Peirce al modelo duda-creencia de Descartes). Pero lo que sí genera dudas es el desacuerdo con nuestros pares. Es curioso que Bilgrami desdeñe la importancia epistemológica de las audiencias a las que se enfrenta el investigador. A Bilgrami parece bastarle decir que, en tanto es el uso cautelar de verdad el que sustenta la premisa 2 de Rorty, y tal uso no tiene efecto en la práctica, entonces la premisa 2 puede ser abandonada desde un punto de vista estrictamente pragmatista. Pero Rorty no se basa en el uso cautelar de verdad para terminar de señalar que la verdad no es el fin de la investigación. Si el fin de la investigación es, desde Peirce y con Bilgrami, la disolución de la duda, la misma no se alcanza hasta que dejamos de ser desafiados por nuestros pares. Una vez hallado el consenso y, en consecuencia, una vez hemos incorporado nuevas creencias a nuestra imagen del mundo, entonces tendremos nuevos candidatos para alabar con el predicado *verdadero*. Pero, ¿sería posible despejar nuestras dudas y tomar a una creencia como verdadera cuando nuestra comunidad no consensúa con nosotros al respecto, no apelando al uso cautelar de verdad sino a detalles relevantes en dicho contexto para la aceptación de dicha creencia? No entiendo cómo uno podría defender dicha posibilidad y permanecer en el pragmatismo. Con lo cual se mantiene la idea de Rorty de que es la justificación, en tanto empresa intersubjetiva a través de la cual alcanzamos el consenso relevante que aquietta nuestras dudas, la meta a la que le podemos dar contenido en la caracterización de la práctica de la investigación. La idea de verdad como fin de la investigación, la idea que sustenta la concepción veroteleologista, sigue sin hacer ninguna diferencia en la práctica, aun adoptando el punto de vista de la primera persona tan caro a Bilgrami.

Por último, quiero referirme a otras posiciones contemporáneas de corte peirceano, como las sostenidas por Apel, Habermas o alguno de los múltiples Putnams, que

también defienden el veroteleologismo, al defender la idea de que hay una diferencia entre la verdad y la justificación que no redunde en diferencias expresivas. Para estas versiones de peírceanismo las prácticas de justificación tienden a la convergencia, y es al polo ideal de convergencia en el acuerdo de las distintas comunidades para el que reservamos la aplicación de la noción de *verdad*, entendida como ideal regulativo de la actividad justificadora. El fin de la investigación no es ya la verdad entendida como correspondencia con la realidad sino como convergencia ideal en la justificación. Las diferencias entre esta posición y la rortiana son sutiles, pero el mismo Rorty se ha empeñado recientemente en señalar cuál es exactamente el punto en disputa.

La idea es que quienes se aferran a la idea de convergencia suponen que hay algo así como una ética de la justificación que no es contextual y que entraña un compromiso universalista, de modo que la práctica de la justificación conduce a un acuerdo racional entre todos los participantes de la comunicación lingüística, siendo así posible decir que hay determinadas comunidades que llevan adelante prácticas de justificación que se alejan notablemente de los presupuestos pragmáticos que todo intercambio comunicativo idealmente supone. La convergencia es posible en tanto todo usuario del lenguaje puede finalmente asumir tales presupuestos y embarcarse entonces en intercambios comunicativos libres de dominación. El acuerdo intersubjetivo y, por tanto, la verdad, guían la práctica de la justificación y están idealmente garantizados en la medida en que pueden ponerse en acto prácticas que no incurran en autocontradicción performativa. “Ellos –dice Rorty, refiriéndose a los teóricos de la convergencia– ven el deseo de verdad, construido como el deseo de hacer proclamas de validez universal, como el deseo de justificación universal” (2000, p. 17).

Frente a este tipo de sugerencias, la respuesta de Rorty es inmediata. No se entiende qué se quiere decir con presupuestos universales de las prácticas de justificación, y es falso que a la base de nuestra práctica de justificación haya un deseo de acuerdo universal. “Davidson y yo –agrega Rorty– no solemos apelar a la idea de que toda acción comunicativa contiene un llamado a la validez universal, porque esta así llamada “presuposición” no nos parece que desempeña papel alguno en la explicación de la conducta lingüística” (*ibid.*, pp. 16-17). A su vez, la práctica de la justificación requiere que logremos acuerdo dentro de nuestra comunidad, pero de ninguna manera precisa que el acuerdo sea universal, integrando a todo potencial interlocutor.

No hay uso del lenguaje sin justificación; no hay habilidad para tener creencias sin habilidad para argumentar acerca de qué creencias tener. Pero decir esto no

es decir que la habilidad para usar el lenguaje, para tener creencias y deseos, supone un deseo de justificar las propias creencias a todos los organismos usuarios del lenguaje que nos encontremos. No todo usuario del lenguaje que se nos interponga en el camino será tratado como miembro de una audiencia competente. Por el contrario, los seres humanos usualmente se dividen en mutuamente suspicaces comunidades de justificación (*ibid.*, p. 15).

El uso del lenguaje requiere la práctica de la justificación, pero no un tipo particular de comunidad lingüística. Nuestra comunidad de justificación democrática asume el ideal de la convergencia, de la maximización del acuerdo, pero éste es un ideal etnocéntrico que no nos debe llevar a ver a otras comunidades, comunidades exclusivistas, por ejemplo, como queriendo jugar el juego de la justificación sin jugarlo. Incluso, diría Rorty, nuestro ideal democrático de convergencia no es siquiera tan fuerte como para comprometernos con la idea de que todo interlocutor es uno frente al cual debo justificar mis creencias. Muchos de los interlocutores, de los usuarios del lenguaje que se nos ponen en el camino, no son tratados como pares en la argumentación, sino como miembros de comunidades a los que tengo que persuadir por medios más sentimentales que argumentativos, pues no hay comunidad argumentativa posible entre ellos y nosotros. Los niños de cuatro años, los racistas y los nazis no son tratados como pares sino como sujetos del lenguaje a los que hay que educar, educación en la que la argumentación suele volverse inútil, haciéndonos imprescindible recurrir a la lectura de cuentos, ya sea *Pulgarcito*, *La cabaña del tío Tom* o *El diario de Anna Frank*.

Pensar que hay algo así como presupuestos universales de la práctica justificadora, presupuestos que conducen idealmente al acuerdo intercomunitario, es otra forma de salirse del etnocentrismo al que conduce inevitablemente la apuesta antirrepresentacionista, asumiendo la cuestionable idea de que la universalidad no debe ser creada, sino que estuvo allí desde siempre a la espera de que nos acomodáramos a ella⁹.

9. Agradezco a los participantes del VII Coloquio Iberoamericano de Filosofía (Bogotá, 2002) por la generosidad en sus comentarios y críticas a una versión previa de este trabajo. También quiero manifestar mi gratitud para con los miembros del denominado Grupo de Acción Filosófica (GAF), en cuyas reuniones en el Instituto de Filosofía de la UBA pude presentar y mejorar la mayoría de las ideas que aquí defiendo.

Bibliografía

- Allen, B. (2000): "What was Epistemology?", en: R. Brandom (ed.), *Rorty and His Critics*, Malden/Oxford: Blackwell.
- Bilgrami, A. (2000), "Is Truth a Goal of Inquiry?: Rorty and Davidson on Truth", en: R. Brandom (ed.), *Rorty and His Critics*, Malden/Oxford: Blackwell.
- Brandom, Robert (1994): *Making it Explicit*, Cambridge, Harvard U.P.
- Haack, Susan (1993): *Evidence and Inquiry*, Oxford-Cambridge: Blackwell.
- Peirce, Charles S. (1931-1958): "The Fixation of Belief", en: Hartshorne, Weiss, Burks (eds.), *Collected Papers of Charles Sanders Peirce*, vol. V, Cambridge: Harvard U.P.
- Penelas, F. (2001)(a): "Deflacionismo y conversacionalismo", en: R. Caracciolo y Diego Letzen (eds.), *Epistemología e Historia de la Ciencia*, Volumen 7, Córdoba, Facultad de Humanidades de la Universidad Nacional de Córdoba.
- Penelas, F. (2001)(b): "Antirrepresentacionalismo y realismo interno: un debate en torno a la verdad y la justificación", en G. Hurtado (ed.), *Subjetividad, representación y realidad*, Puebla, Benemérita Universidad Autónoma de Puebla.
- Penelas, F. (2003): "La justificación como hecho social", en: *Dianoia*, Vol. XLVIII, No. 51.
- Rorty, Richard (1991): "Pragmatism, Davidson and Truth", en: R. Rorty: *Objectivity, Relativism and Truth*, Cambridge: Cambridge U.P.
- Rorty, Richard (1995): "Response to Susan Haack", en: H. Saatkamp Jr. (ed.), *Rorty and Pragmatism*, Nashville: Vanderbilt University Press.
- Rorty, Richard (1998): "Is Truth a Goal of Inquiry? Davidson vs. Wright", en: R. Rorty: *Truth and Progress*, Cambridge: Cambridge U.P.
- Rorty, Richard (2000): "Universality and Truth", en: R. Brandom (ed.), *Rorty and His Critics*, Malden/Oxford: Blackwell.
- Williams, Michael (1999): *Groundless Belief*, Princeton: Princeton University Press.

**La explicación científica:
causalidad, unificación
y subsunción teórica***

Universitat Rovira i Virgili (Tarragona)
y Logos Research Group on Language,
Logic and Cognition (Barcelona).

1. Explicación y explicación científica

La palabra “explicar” tiene diferentes usos en castellano, al menos los tres siguientes: “narrar” un acontecimiento, como en “María me explicó su viaje a la India”; “elucidar” o “aclarar” una palabra, concepto o idea, como en “Juan me explicó el concepto de espín” (se corresponde con el inglés *explication*); “dar razón” de por qué sucedió algo, como en “el coche se salió de la calzada porque había arena en la carretera” (se corresponde con el inglés *explanation*). Nosotros estamos interesados aquí por el último de estos usos.

En una primera aproximación intuitiva, en este uso una explicación se puede considerar una respuesta a una “pregunta por qué” (*why-question*): “¿Por qué se salió el coche de la calzada? Porque había arena en la carretera”. “¿Por qué se hundió el Titanic? Porque chocó con un iceberg”. En general: “¿Por qué α ? Porque γ ”, siendo α el *tema* de la pregunta y γ la respuesta que constituye una explicación. Como ha enfatizado Van Fraassen (1977), esta primera aproximación es incompleta, pues una “pregunta por qué” no queda especificada sólo por el tema. Por ejemplo, la pregunta “¿Por qué fue Juan a la fiesta?” no es completa en tanto que demanda de explicación, pues puede estar requiriendo explicación de cosas muy diferentes, y la adecuación de la respuesta en tanto que explicación depende de cuál de tales cosas se requiera. En cierto contexto, se puede estar requiri-

* Este trabajo forma parte de los proyectos de investigación BFF2002-04454-C10-05, Ministerio Español de Ciencia y Tecnología, y BFF2002-10164-E, European Science Foundation. Agradezco a Jesús Zamora, Ulises Moulines y al resto de participantes del III Encuentro Internacional de Metateoría Estructuralista (Granada, marzo, 2002), así como a los participantes del VII Coloquio Iberoamericano de Filosofía (Bogotá, septiembre, 2002) por sus comentarios críticos a versiones anteriores de este trabajo.

riendo explicación del hecho de que fuera Juan a la fiesta *en lugar de Pedro o María*, es decir: ¿Por qué fue *Juan* a la fiesta? En este caso, una respuesta como “porque quería ir al cine o al teatro o a la fiesta y no encontró entradas ni para el cine ni para el teatro” no constituye una explicación satisfactoria, mientras que una respuesta como “porque tenía que ir uno de los tres pero ni Pedro ni María podían, y él sí” sí constituiría una buena explicación. En otro contexto, sin embargo, esa “misma” pregunta puede estar requiriendo explicación del hecho de que Juan fuera a la fiesta *en lugar de al cine o al teatro*, esto es: ¿Por qué fue Juan *a la fiesta*? En este caso, la primera respuesta constituye una buena explicación pero la segunda no. Así, en general, una “pregunta por qué” no queda plenamente determinada especificando sólo el tema α , hay que especificar también una *clase de contraste*, el conjunto de alternativas en contraste con las cuales la ocurrencia de α requiere explicación: “¿Por qué α (en lugar de β_1, β_2, \dots)?” Ahora sí, una vez especificada completamente la “pregunta por qué” con el tema y la clase de contraste, se puede juzgar si la respuesta “porque γ ” constituye una explicación satisfactoria. El lector objetará que en muchas ocasiones no hace falta especificar la clase de contraste, pues en el contexto “está claro qué se pregunta”. Eso es cierto, pero no implica que en esas ocasiones no haya alguna clase de contraste; sólo indica que en esos casos el contexto determina implícitamente la clase de contraste y que no hace falta “hacerla explícita”. De hecho, en muchos de los ejemplos que veremos se va a proceder sin especificar la clase de contraste, justamente porque el contexto la determinará implícitamente.

Nuestro objetivo aquí es analizar el tipo de explicación característica de los contextos científicos. Nos vamos a centrar, pues, principalmente en la explicación científica, aun reconociendo que, *prima facie*, debe haber un núcleo conceptual común a la explicación científica y a la ordinaria. Entre una explicación claramente no científica, como “el coche se salió de la calzada porque había arena en la carretera” y otra claramente científica, como “la luz se polariza porque está formada por corpúsculos”, puede haber todo un continuo de casos intermedios, por ejemplo:

- “El coche se salió de la calzada porque había arena en la carretera”
- “Juan no fue a la fiesta de Rosa porque está enfadado con ella”
- “Las cañerías se rompieron porque se congeló el agua”
- “Pedro tiene cáncer de pulmón porque fumaba cinco paquetes diarios”
- “La bajada brusca de la presión atmosférica explica la tormenta”
- “La fuga radiactiva explica las malformaciones de la población cercana a Chernobil”

“La presencia de Neptuno explica la anomalía en la órbita de Urano”

“La luz se polariza porque está formada por corpúsculos”

Parece en principio obvio que hay un núcleo conceptual común a todos estos casos, y si ello es así entonces es necesario aceptar que hay un núcleo conceptual compartido por las explicaciones científicas y las ordinarias. Es pues un *desideratum* de un análisis satisfactorio de la explicación científica que en principio pueda ser generalizable a la explicación precientífica ordinaria, o que de lo contrario dé buenas razones que justifique su especificidad¹.

En cuanto a la explicación científica, antes de entrar en el análisis propiamente dicho conviene distinguir ya preanalíticamente diferentes tipos o casos. En primer lugar tenemos *explicaciones “deterministas” de hechos particulares* (EDP), por ejemplo, “la presencia de un asteroide explica la desviación del cohete”. En estos casos, el suceso de que se da cuenta (el desvío de cierto cohete) es un suceso particular, esto es, que a cierto objeto particular le sucede cierto fenómeno, y además la explicación es tal que *no deja lugar a que no ocurra*, “determina totalmente” la ocurrencia del hecho particular por explicar. Pero no siempre que explicamos un hecho particular es así. A veces la explicación de un hecho particular es “indeterminista”, esto es, aunque explica que haya ocurrido, la explicación es compatible con que el hecho por explicar en cuestión no hubiese ocurrido. Por ejemplo: “Que los ocho bisabuelos de Rosa hayan sido rubios explica que Rosa también lo sea”. En estos casos decimos que estamos ante *una explicación indeterminista de hechos particulares* (EIP). Por último, a veces el hecho por explicar no es un hecho particular –por ejemplo que a un objeto concreto le sucede algo–, sino un hecho general, esto es, que a objetos de cierto tipo les suceden cosas de cierto tipo; por ejemplo el hecho de que los planetas se muevan de cierto modo (leyes de Kepler). Estos hechos también pueden requerir, y obtener, explicación, por ejemplo, en el caso de los planetas: “La mecánica gravitatoria celeste explica la órbita de los planetas”. Llamaremos a estas explicaciones *explicaciones de hechos generales* (EG)².

Vamos a ver ahora el primero de los análisis de la explicación científica. Todo análisis ha de especificar: (a) qué es el *explanandum* (Exm), esto es, el hecho que requiere explicación; (b) qué es el *explanans* (Exs), esto es, el hecho o hechos que proporcionan la explicación, y (c) qué es la *relación explicativa* (EXP), esto es, la relación tal que en virtud de que se da entre Exs y Exm podemos decir que el pri-

1. Como veremos, éste será el principal reto para el análisis que defenderemos.
2. En el siguiente apartado veremos por qué no se distingue aquí entre explicaciones deterministas e indeterministas, pero avanzamos ya que en este caso todas son deterministas, incluso si el hecho general por explicar es él mismo indeterminista.

mero explica el segundo. Todo análisis que elucide estos elementos ha de satisfacer, *prima facie*, al menos las siguientes tres condiciones de adecuación:

- (i) *Adecuación extensional*: el análisis ha de ser compatible con nuestras intuiciones más firmes sobre casos de explicación: no puede incluir casos claros o paradigmáticos de no explicación ni excluir casos claros o paradigmáticos de explicación; sobre los casos no claros, el análisis puede tener una función regulativa.
- (ii) *Generalidad y especificidad*: el análisis ha de elucidar a la vez qué tienen en común los casos EDP, EIG y EG, y que tiene de diferente o específico cada uno, esto es, en qué se asemejan y en qué difieren estos tres tipos de explicación.
- (iii) *Elucidación filosófica*: el análisis ha de ser filosóficamente “clarificador”, no puede elucidar el concepto de explicación utilizando como primitivos otros conceptos más oscuros; no puede elucidar lo poco claro mediante lo menos claro.

2. Análisis inferencialista nomológico

El primer análisis suficientemente refinado del concepto de explicación científica se debe fundamentalmente a Hempel, quien establece sus líneas generales en un artículo seminal escrito con Oppenheim y lo desarrolla después en otros trabajos en solitario (*cf.* Hempel y Oppenheim, 1948; Hempel, 1965). Este análisis se ha calificado a menudo de “modelo de cobertura legal” (*covering law model*), pero preferimos aquí denominarlo “inferencialista nomológico” para distinguirlo de otras propuestas posteriores que también exigen la cobertura legal.

El análisis inferencialista nomológico surge de la combinación de dos intuiciones básicas, a las que se añade una cualificación. Las dos intuiciones son las siguientes. Según la primera, en una explicación el Exs explica el Exm en el sentido de que lo hace “esperable”. El Exm es en cierto sentido relevante “inesperado” o “sorprendente”, y ofrecer una explicación del mismo es ofrecer otro fenómeno en relación con el cual el Exm ya no es inesperado³. La segunda in-

3. Obviamente, el sentido en que aquí se dice que el Exm es inesperado o sorprendente no implica que sea un fenómeno extraño o de ocurrencia poco habitual. El Exm puede ser a veces un fenómeno extraño o poco usual, como el desvío de un cohete, pero muchas otras veces es un fenómeno familiar, como la salida del sol cada mañana. Aunque en un sentido el segundo fenómeno es “esperable” y el primero no, en otro sentido, el relevante para la explicación, ambos fenómenos son “inesperados” *mientras no se ofrece una explicación*: incluso en los casos en que se sabe que ocurre frecuentemente, sigue siendo sorprendente.

tuición establece que un hecho X es esperable-relativamente-a otro hecho Y si X *se infiere* de Y. La combinación de estas dos intuiciones proporciona el esquema básico de este análisis de la explicación como inferencia (EI- en espera de la cualificación que introduciremos más adelante).

EI: $Exs \text{ EXP } Exs \equiv$ hay una inferencia válida con Exs de premisas y Exm de conclusión

Para este análisis, por tanto, la relación EXP de explicación en virtud de la cual el Exs explica el Exm es la relación de *inferencia lógica*. O mejor, *alguna* relación de inferencia lógica, pues como veremos en seguida, puede haber varias, y diferentes tipos de explicación requieren diferentes tipos de inferencia lógica. Este esquema es el que motiva el calificativo de *inferencialista* que hemos dado al análisis. Sin embargo, como ya hemos anunciado, EI no expresa satisfactoriamente el análisis inferencialista nomológico, pues dicho análisis contiene una cualificación que EI no recoge, la cualificación connotada por “nomológico”. La idea básica que se encuentra tras esta cualificación es que toda explicación ha de ser *general*, ha de tener algo de repetible o regular, y que además esta regularidad involucrada en la explicación no puede ser meramente *accidental* sino que ha de ser una regularidad *nomológica*, una ley natural. Esta cualificación queda recogida en el siguiente esquema, que expresa de forma completa todos los elementos involucrados en el análisis de la explicación como inferencia nomológica (EIN):

EIN: $Exs \text{ EXP } Exs \equiv$ hay una inferencia válida con Exs de premisas y Exm de conclusión y el Exs contiene al menos esencialmente⁴ una regularidad nomológica y todas las regularidades que contiene esencialmente son nomológicas.

No vamos a justificar ahora en general esta cualificación. Vamos a ver su justificación para el caso concreto de las explicaciones deterministas de hechos particulares (EDP) y veremos entonces que es naturalmente generalizable a otros casos.

2.1 Explicación nomológico-deductiva de hechos particulares

Habíamos caracterizado preanalíticamente las explicaciones indeterministas de hechos particulares como aquellas en las que el Exm es un acaecimiento particular y que son tales, que la explicación “no deja lugar para la no ocurrencia” del Exm. El caso que habíamos mencionado como ejemplo es el de la explicación del desvío de la trayectoria de cierto cohete *a* por el paso de cierto asteroide *b* de tal masa

4. “Esencialmente” significa que es esencial para la inferencia, que es una premisa necesaria para la inferencia (pues siempre es posible añadir premisas adicionales superfluas, y sobre ellas el análisis no dice nada).

a tal distancia de a . Dado el paso de dicho asteroide (y si ningún otro cuerpo interviene), el cohete no puede no alterar su trayectoria como lo hace. La idea de Hempel es que la relación inferencial en lógica análoga a este “no dejar lugar para que [lo inferido] no ocurra” es la relación de *inferencia deductiva*, de modo que en las explicaciones deterministas (de hechos particulares) la relación EXP es la relación de deducción: el Exs explica el Exm en el sentido de que el Exm se *infiere deductivamente* del Exs. Ello quiere decir que la explicación dada del desvío del cohete constituye implícitamente un argumento deductivo válido que tiene al paso del asteroide entre sus premisas y al desvío del cohete como conclusión⁵. El argumento tiene el paso del asteroide “como una de sus premisas” y no “como [todas] sus premisas”, porque para que el argumento sea deductivamente válido no puede tener como única premisa el paso del asteroide. En efecto, el argumento (connoto que un argumento es deductivo mediante una raya continua)

El asteroide b de masa m pasa a la distancia d del cohete a

El cohete a desvía su trayectoria de modo t
 es un argumento deductivamente inválido. Si la explicación nos parece satisfactoria es porque además del hecho explicitado del paso del asteroide consideramos que el Exs contiene implícitamente otro hecho, un hecho que “conecta” el paso del asteroide con el desvío del cohete. Y aquí es donde se justifica la cualificación a que hicimos mención sobre regularidades nómicas. Ese hecho que conecta el paso del asteroide con el desvío del cohete no puede ser cualquiera que constituya una premisa adicional que convierta al argumento en deductivamente válido, no puede ser, por ejemplo, simplemente “si el asteroide b de masa m pasa a la distancia d del cohete a entonces el cohete a desvía su trayectoria de modo t ”. Intuitivamente –y Hempel así lo reconoce–, el argumento con esas dos premisas, aunque deductivamente válido, no constituiría una explicación satisfactoria del desvío del cohete. El motivo es que la explicación carece de *generalidad*; el hecho adicional que conecta el paso del asteroide con el desvío del cohete, que conecta esos dos hechos particulares, es también un hecho, aunque condicional, particular que involucra sólo a esos objetos: que si ese asteroide se comporta así entonces ese co-

5. A lo largo del trabajo, y mientras no sea necesario proceder de otra manera, me referiré a las premisas y conclusiones de argumentos indistintamente como enunciados (expresiones lingüísticas) o como hechos o proposiciones (lo que las expresiones lingüísticas expresan). Para ciertos fines es conveniente distinguir ambos casos, pero aquí no lo haré, salvo si es necesario.

hete se comporta así. Eso es intuitivamente insatisfactorio en tanto que explicación. Para que la explicación sea satisfactoria, el hecho que conecta el Exm particular con el resto de hechos particulares del Exs debe ser un hecho general, una *regularidad*; por ejemplo, “siempre que un cuerpo de tal masa pasa a tal distancia de otro le altera su trayectoria de tal y cual modo” (sería más exacta como premisa adicional la conjunción de la ley de gravitación con otras leyes dinámicas, pero valga esta simplificación informal a los actuales efectos):

El asteroide b de masa m pasa a la distancia d del cohete a [de masa m']

Siempre que un cuerpo de masa x pasa a distancia y de otro cuerpo de masa x' altera su trayectoria de modo z

El cohete a desvía su trayectoria de modo t

Este argumento, deductivamente válido, sí parece ahora una buena explicación de nuestro Exm, el desvío del cohete. Y si lo parece, es porque ese hecho conector al que hicimos referencia es una regularidad. En realidad, no basta con que sea una regularidad. Como el propio Hempel señala, si la regularidad fuese meramente accidental, tampoco nos parecería una explicación satisfactoria. Las regularidades incluidas en el Exs que conectan los hechos particulares del resto del Exs con el hecho particular Exm no pueden ser meras regularidades accidentales: han de ser *regularidades nómicas*, leyes naturales⁶. Y así sucede con la regularidad sobre cruce de cuerpos y alteraciones de trayectorias referida en el Exs de la explicación del desvío del cohete; por eso nos parece una explicación satisfactoria.

El siguiente esquema resume las explicaciones nomológico-deductivas de hechos particulares (NDP), el análisis inferencialista nomológico aplicado a las explicaciones deterministas de hechos particulares: ofrecer una explicación determinista de un hecho particular es dar una serie de otros hechos particulares (*condiciones antecedentes*) y de leyes de los que conjuntamente se deduzca el hecho por explicar (connoto los hechos particulares mediante minúsculas y los generales mediante mayúsculas):

6. Es una cuestión sustantiva y difícil especificar cuál es la diferencia entre una regularidad accidental (por ejemplo, “todas las monedas de mi bolsillo ahora son amarillas”) y una regularidad nómica (por ejemplo, “todas las monedas de oro son amarillas”), pero no podemos entrar aquí en ello. Basta con la idea intuitiva y con la constatación de que este análisis presupone esta distinción. Para una exposición de las diferentes elucidaciones de esta diferencia, cf. Díez y Moulines, 1997, cap. 5.

NDP	$\frac{c_1, \dots, c_n}{L_1, \dots, L_j}$	argumento deductivo válido y las L_i (esenciales) son regularidades nómicas
	e	

2.2 Explicación nomológico-inductiva de hechos particulares

Hemos caracterizado preanalíticamente las explicaciones indeterministas de hechos particulares como aquéllas en las que el Exm es un acaecimiento particular y que son tales que la explicación, aun siendo satisfactoria, sería compatible con la no ocurrencia del Exm. El caso mencionado como ejemplo es el de la explicación del hecho de que Rosa es rubia apelando a que sus ocho bisabuelos son a su vez rubios; análogamente sucede en la explicación de la enfermedad respiratoria de Pedro apelando a que ha consumido 50 cigarros diarios durante 40 años. Acabamos de ver que en las explicaciones deterministas, la idea de Hempel es que la relación inferencial en lógica análoga a “no dejar lugar para que [lo inferido] no ocurra”, es la relación de inferencia deductiva. Ahora, en los casos de explicación indeterminista, su idea es que la relación inferencial en lógica correspondiente al carácter indeterminista, de estas explicaciones es la relación de *inferencia inductiva*. En las explicaciones indeterministas la relación EXP es la inferencia inductiva: el Exs explica el Exm en el sentido de que el Exm se *infiere inductivamente* del Exs. Ello quiere decir que la explicación dada del hecho de que Rosa es rubia constituye implícitamente un argumento inductivo válido (o fuerte) que tiene al hecho de que sus ocho bisabuelos son rubios como una de sus premisas y el hecho de que Rosa es rubia como conclusión. Análogamente al caso determinista, para que el argumento sea inductivamente válido no puede tener como única premisa ese hecho. Si la explicación nos parece satisfactoria es porque además del hecho explicitado consideramos que el Exs contiene implícitamente otro hecho, un hecho que “conecta” el color del pelo de los bisabuelos de Rosa con el color del pelo de Rosa. Y por motivos análogos a los expuestos para las explicaciones deterministas (esto es, que toda explicación contiene “generalidad”), tampoco aquí este hecho conector puede ser otro hecho particular: ha de ser un hecho general, una regularidad. Y, por idénticos motivos que antes, no puede ser una regularidad meramente accidental: ha de ser una regularidad *nómica*, una ley natural; por ejemplo, en nuestro caso, la regularidad nómica “la probabilidad de ser rubio si los ocho bisabuelos lo son es cercana a 1”. Así, el argumento (connoto el carácter inductivo con una línea discontinua)

Los ocho bisabuelos de Rosa son rubios

La probabilidad de ser rubio si los ocho bisabuelos lo son es cercana a 1

Rosa es rubia

sería, según la propuesta de Hempel, una buena explicación indeterminista de nuestro Exm. Puesto que estos argumentos son inductivos, la regularidad nómica que permite la inferencia inductiva será una regularidad probabilista o estadística. El siguiente esquema resume las explicaciones nomológico-inductivas de hechos particulares (NDP): ofrecer una explicación indeterminista de un hecho particular es dar una serie de otros hechos particulares (*condiciones antecedentes*) y de leyes (entre ellas al menos una probabilista) de los que conjuntamente se infiera inductivamente el hecho por explicar.

NIP	$\frac{C_1, \dots, C_n}{L_1, \dots, L_j}$	argumento inductivo válido y las L_i (esenciales) son regularidades nómicas, al menos una de las cuales es probabilista/estadística
	e	

2.3 *Explicación nomológica-deductiva de hechos generales*

Habíamos caracterizado las explicaciones de hechos generales como aquéllas en las que el Exm es un hecho general, esto es, una regularidad: a objetos de cierto tipo les suceden fenómenos de tal tipo. Los hechos generales también pueden requerir explicación, y ésta consta, parece, de otros hechos “más generales”. Por ejemplo, explicamos que los planetas se muevan de cierta manera (leyes de Kepler), apelando a la mecánica gravitatoria celeste; o que los rasgos fenotípicos se transmitan de cierta manera (leyes de Mendel) apelando a las leyes de la genética molecular. Según Hempel, también en estos casos, como en las explicaciones de hechos particulares, la explicación satisfactoria constituye implícitamente un argumento válido, la relación EXP es ahora también una relación de inferencia lógica. Pero, ¿qué relación de inferencia? Hemos visto que en la explicación determinista de hechos particulares la relación EXP era la de inferencia deductiva, y en la explicación indeterminista de hechos particulares EXP es la relación de inferencia inductiva. ¿Hay una diferencia semejante en las explicaciones EG, de hechos generales? El que no hayamos distinguido preanalíticamente aquí dos casos sugiere que la respuesta es negativa. Y efectivamente así lo entiende también Hempel. Es cierto que podemos distinguir aquí Exm que son regularidades absolutas o deterministas, como las leyes de Kepler, de Exm que son regularidades probabilistas o indeterministas, como las leyes de Mendel. Pero de ahí no se sigue que a esa diferencia en la naturaleza del Exm le corresponda otra en la naturaleza de la relación explicativa. En ambos casos, EXP es la inferencia deductiva; el Exm se infiere del Exs *deduciéndolo* del Exs. La diferencia

entre ambos casos no está en la relación explica-tiva sino en el *explanans*: la explicación de hechos generales deterministas consiste en *deducir* tales hechos de otros hechos generales *también deterministas* (así, por ejemplo, es como explicamos las leyes de Kepler, deterministas, deduciéndolas de las leyes mecánicas gravitatorias, también deterministas); y la explicación de hechos generales *indeterministas* consiste también en *deducir* tales hechos de otros hechos generales, sólo que ahora esos hechos que constituyen el Exs serán (al menos uno) *indeterministas* (así, por ejemplo, explicamos las leyes de Mendel, indeterministas, deduciéndolas de otras de la genética molecular, también indeterministas). El siguiente esquema resume los rasgos principales de las explicaciones nomológico-deductivas de hechos generales (NDG):

$$\text{NDG} \quad \frac{L_1, \dots, L_j}{E} \quad \text{argumento deductivo válido}$$

En algunos casos el esquema sería un poco más complejo, pues a veces la explicación es “idealizada”; el Exs contiene condiciones de *contorno* necesarias para la inferencia (por ejemplo, en la deducción de las leyes clásicas a partir de las relativistas). Pero ello no afecta a la idea central, que basta para nuestros actuales fines.

2.4 Problemas del análisis inferencialista

El análisis inferencialista satisface bastante bien dos de los tres *desiderata* que establecimos. Por un lado, elucida claramente qué es lo que los tres tipos de explicación tienen en común, esto es, todos son inferencias lógicas válidas con el Exm de conclusión cuyas premisas, el Exs, contienen (al menos) regularidades nomológicas; y también qué es lo que tienen de diferente, esto es, lo que especifican los esquemas NDP, NIP y NDG. Por otro, parece satisfacer bien la exigencia de clarificación filosófica, pues elucida un concepto oscuro, el de *explicación*, mediante otros en principio más claros, el de *inferencia* y el de *ley*⁷.

7. Digo que “parece” satisfacerlo porque para algunos críticos el concepto de ley es tan oscuro como el de explicación (de hecho Hempel mismo intentó una elucidación humeana de dicho concepto, sin éxito). Pero en cualquier caso, podemos conceder que el concepto de ley tiene dificultades independientes y que precisa elucidación por motivos independientes y comunes a los diferentes análisis de la noción de explicación, de modo que por ese lado no está peor que los rivales. Por otro lado, la noción de argumento válido tampoco es totalmente clara, pero sí algo más que la de explicación (salvo quizás la de validez inductiva, pero de nuevo esto es un problema general en el que Hempel no está en desventaja respecto de los rivales).

Sin embargo, este análisis tiene dificultades graves con el otro *desideratum*, la adecuación extensional. Hempel da condiciones necesarias y suficientes para el concepto de explicación, esto es, establece ciertas condiciones tales que si se cumplen estamos ante una explicación satisfactoria y si no se cumplen no estamos ante una explicación satisfactoria. Pues bien, los problemas consisten en que tales condiciones no parecen, bien miradas las cosas, ni necesarias ni suficientes. Hay cosas que las cumplen de las que no diríamos que son buenas explicaciones, y hay cosas de las que diríamos que son buenas explicaciones y no las cumplen. Veámoslo (nos centraremos en la explicación de hechos particulares).

Contraejemplos a la suficiencia de NDP/ NIP

Hay casos de argumentos válidos que cumplen las condiciones de NDP/NIP pero que intuitivamente no diríamos, los científicos no dicen, que las premisas explican el *explanandum*. Veamos algunos:

- (a) *Temporalidad*. Podemos deducir la ocurrencia de un eclipse a partir de la posición anterior de los astros y de ciertas leyes mecánicas y ópticas, y ello constituye, parece, una explicación satisfactoria del eclipse. Pero sucede que también podemos deducir válidamente el eclipse a partir de esas mismas leyes y la posición de los astros *después* del eclipse. Esta segunda inferencia satisface tanto como la primera las condiciones NDP, y sin embargo no es un caso de explicación: no parece aceptable decir que la posición de los astros después del eclipse (más ciertas leyes) *explica* el eclipse⁸.
- (b) *Simetría*. Podemos deducir la longitud de una sombra de un mástil a partir de la longitud del mástil, la inclinación del sol y ciertas leyes ópticas, y ello constituye, parece, una explicación satisfactoria de la longitud de la sombra. Pero también podemos deducir válidamente la longitud del mástil a partir de la de la sombra, y la inclinación del sol y las mismas leyes. Esta segunda inferencia satisface tanto como la primera las condiciones NDP y, sin embargo, no parece aceptable decir que la longitud de la sombra *explica* que el mástil mida lo que mide.
- (c) *Efectos de causa común*. Podemos deducir (o inducir) la ocurrencia de una tormenta a partir del descenso brusco de un barómetro en las cercanías y la regularidad no accidental de que siempre (o casi siempre) que baja bruscamente el barómetro sobreviene una tormenta. Esta inferencia satisface las

8. No se cuestiona que lo pueda “predecir”, o mejor retrodecir. Eso es simplemente que la inferencia es válida. Lo que está en cuestión es si eso basta para considerarlo una *explicación*.

condiciones NDP (o alternativamente NIP), pero no diríamos que el descenso brusco del barómetro *explica* la tormenta. Más bien decimos que es el descenso brusco de la presión atmosférica en los alrededores el que explica la ocurrencia de la tormenta, y por cierto también el descenso brusco del barómetro.

- (d) *Irrelevancia*. Podemos deducir el no embarazo de Juan a partir del hecho de que ha tomado pastillas anticonceptivas y de la regularidad no accidental de que nadie que toma pastillas anticonceptivas queda embarazado. Esta inferencia satisface las condiciones NDP, pero no diríamos que la ingestión de pastillas *explica* el no embarazo de Juan. Diríamos que ese Exm se explica por el hecho de que Juan es varón (y la regularidad no accidental de que los varones no quedan embarazados), y que la ingestión de pastillas es explicativamente irrelevante.

El tercer caso puede ser también un caso de contraejemplo a NIP (si las regularidades involucradas son probabilistas, como parece), y es fácil buscar contraejemplos a NIP similares a cada uno de los otros casos. Así, parece claro que hay casos en que *sí* se satisfacen NDP/NIP pero que *no* aceptamos como explicaciones satisfactorias; no todos los casos de NDP/NIP son buenas explicaciones deterministas/indeterministas. El análisis de Hempel, por tanto, no establece, contra lo que pretende, condiciones suficientes para la explicación. Para descartar esos contraejemplos hay que exigir algo más (o algo diferente) a inferencias de esos tipos.

Contraejemplos a la necesidad de NIP

NIP establece las condiciones tales que, si algo las cumple, entonces puede ser considerado legítimamente –según la propuesta de Hempel–, un caso de explicación indeterminista de un hecho particular. Y, sin embargo, hay casos en los que concordaríamos en que estamos ante una buena explicación indeterminista pero tales que, reconstruidos como argumentos inductivos, resultan ser malos argumentos inductivos, inferencias inductivas inválidas.

El contraejemplo más famoso se refiere a la paréisis, una enfermedad que se desarrolla en aproximadamente el 25% de los casos de enfermos que contraen sífilis y no se tratan adecuadamente con antibióticos. Supongamos que el alcalde tiene paréisis y que ese es nuestro Exm, es decir, queremos una explicación de que contrajese paréisis. Consideremos el siguiente *explanans*: “Porque contraí sífilis, no se trató adecuadamente con antibióticos y el 25% de los que contraen sífilis y no se tratan adecuadamente con antibióticos desarrollan paréisis”. Intuitivamente, parece que esa es una buena explicación. Ante una respuesta de ese tipo, uno no sigue preguntando, la demanda de explicación ha sido satisfecha (o tan satisfecha como con cualquier otra explicación indeterminista satisfactoria). Y, sin embargo,

si reconstruimos esta explicación como un argumento inductivo con el Exs de premisas y el Exs de conclusión resulta un mal argumento inductivo, un argumento inductivo inválido. Así, parece que hay *buenas* explicaciones indeterministas que reconstruidas como argumentos inductivos son *malos* argumentos inductivos, no satisfacen NIP. Por tanto, las condiciones que establece NIP no parecen, contra lo que pretende el análisis de Hempel, condiciones necesarias para una buena explicación indeterminista.

3. Análisis alternativos: causalismo y unificacionismo

Vamos a ver aquí los principales rasgos de las dos propuestas alternativas más importantes presentadas para superar las deficiencias del análisis inferencialista nomológico. Ellas se ocupan mayoritariamente de los contraejemplos (a)-(d) a la suficiencia, que afectan tanto a las explicaciones deterministas como a las indeterministas, pero deben decir algo también respecto de los contraejemplos, como el de la paresis, a la necesidad, aunque afecten sólo a las explicaciones indeterministas. Antes de pasar a ver estas alternativas, mencionaremos muy brevemente la propuesta “técnica” de Salmon respecto a las explicaciones indeterministas, pues en principio es una modificación del análisis de este tipo de explicación que parece básicamente correcta y, por tanto, con la que toda propuesta filosófica más general (que no rechace la legitimidad de las explicaciones indeterministas) debería ser al menos compatible.

3.1 Relevancia estadística

El análisis NIP de las explicaciones indeterministas exige que el Exs constituya las premisas de una inferencia inductiva válida que tiene como conclusión el Exm. Ello equivale a decir que una explicación indeterminista es satisfactoria cuando el Exs *hace muy probable* el Exm, esto es, la probabilidad del Exs relativamente al Exm, $p(\text{Exm}/\text{Exs})$, es muy alta. Hemos visto que el ejemplo de la paresis contradice esta condición, a pesar de ser intuitivamente una buena explicación indeterminista. La idea de Salmon (1984, cap. 2; 1989) es que en las explicaciones indeterministas, la condición para que el Exs explique satisfactoriamente el Exm no es que lo haga *muy probable*, sino sólo que lo haga *más probable*, más probable que si no ocurre el Exs. La relación EXP es en este caso no la de inferencia inductiva, sino la de *relevancia estadística (positiva)*: $p(\text{Exm}/\text{Exs}) > p(\text{Exm}/\neg\text{Exs})$ ⁹. Esta propuesta trae consigo diversas

9. En realidad, Salmon exige sólo relevancia estadística, esto es, que la probabilidad del Exm dada la ocurrencia del Exs sea diferente a la probabilidad dada su no ocurrencia (cf. Salmon, 1989 p. 67).

complicaciones para evitar algunas dificultades, pero no entraremos en ellas. Lo apuntado basta para los actuales fines.

3.2 *Explicación y causalidad*

La idea central del modelo causal de explicación es que “explicar un acontecimiento es proporcionar información acerca de su historia causal” (Lewis 1986, p. 217). En la explicación de un hecho el Exs no tiene por qué asegurar la ocurrencia del Exm, tampoco hacerlo altamente probable, o más probable. Todo ello está vinculado a la idea hempeliana de esperabilidad, pero la noción de explicación no tiene que ver, al menos no directamente según el causalista, con esa idea. Explicar un hecho no es mostrar que es (totalmente, mucho, o tan sólo más) esperable, es proporcionar *información causal* sobre su ocurrencia. Quizás a veces, o incluso en muchas ocasiones, la explicación confiere cierta esperabilidad al Exm, pero ello es así sólo derivativamente, consecuencia de que a veces la información sobre la historia causal, que es el objeto básico de la explicación, tiene ese efecto.

Que explicar consista conceptualmente en proporcionar información causal no implica que una explicación deba ofrecer *toda* la información causal del Exm, o que sea tanto mejor cuanto más información causal proporcione. La historia causal completa de un acaecimiento incluye multitud de otros acaecimientos, y la mención a todos ellos no sólo no es necesaria sino que sería incompatible con una buena explicación. Cada hecho particular tiene muchos otros hechos antecedentes como causas; en cada momento del pasado de un hecho hay una multiplicidad de hechos que son *causas parciales* del mismo y que conjuntamente constituyen su *causa total* en ese momento. Además, dicha multiplicidad causal suele variar en cada momento del pasado de un suceso. Así, *la historia causal completa* de un hecho recoge el conjunto de todas las causas parciales antecedentes, el conjunto de causas totales en cada momento. Pues bien, una explicación no requiere informar sobre toda la historia causal; ni siquiera, generalmente, sobre la causa total en un momento antecedente dado. En general se exige sólo información sobre *algunos* factores causales. Cuáles son esos factores es algo que depende de cada contexto explicativo, el contexto determina qué antecedentes causales se consideran relevantes o destacados a efectos explicativos en esa ocasión. La relación EXP es, pues, en este análisis la relación de *relevancia causal*. Que la relación de relevancia causal sea pragmática y sensible al contexto no significa que no sea objetiva. El contexto determina cuáles son los factores explicativamente relevantes, *de entre los que contiene la historia causal completa*, que es plenamente objetiva e independiente del contexto. Tenemos pues el siguiente esquema que resume el análisis causalista:

CAUS Exs EXP Exm (en un contexto C) syss Exs proporciona los antecedentes causales relevantes-en-C de entre los contenidos en la cadena causal completa de Exm

Éste es el núcleo del análisis causalista¹⁰. Es fácil ver que este análisis resuelve satisfactoriamente los contraejemplos (a la suficiencia de NDP/NIP) anteriores. De hecho, según sus defensores, dichos contraejemplos muestran justamente que las explicaciones satisfactorias consisten en proporcionar información causal, pues lo que falla en los contraejemplos es justamente que en esas inferencias las premisas no son parte de la historia causal de la conclusión. Por ejemplo, en los casos de prioridad temporal, las “explicaciones” en las que el Exs es posterior al Exm (eclipse actual “explicado” por posiciones futuras) no son válidas, pues el *explanans* debe ser causa del *explanandum*, y las causas preceden a sus efectos. Análogamente con los casos de simetría (la sombra y el mástil), causa común (el barómetro y la tormenta) e irrelevancia (la ingestión de pastillas anticonceptivas y el no embarazo de Juan). En todos ellos, cuando la respuesta parece una buena explicación es porque incluye información causal sobre el Exm, y cuando no lo parece es porque no la incluye.

Veamos cómo le va al modelo causalista con los *desiderata* con que evaluamos el modelo hempeliano. En cuanto a la adecuación extensional, parece que resuelve al menos los contraejemplos que se presentaban a la suficiencia del análisis hempeliano. Se ha argumentado, sin embargo, que el análisis causalista carece de sus propios contraejemplos pues, se aduce, hay explicaciones cuya naturaleza causal es, cuando menos, discutible. Estos contraejemplos son de dos tipos. En primer lugar, las explicaciones de hechos particulares indeterministas (como la enfermedad pulmonar de Juan, el atravesar un electrón una barrera de potencial o la paresis del alcalde). Las explicaciones de estos hechos se pueden considerar causales sólo si se acepta causalidad probabilista, lo que algunos rechazan¹¹. El segundo tipo de contraejemplos proviene

10. Nótese que aparentemente este esquema no hace referencia a leyes, pero sólo aparentemente. No hay referencia *explícita* a las leyes pero sí *implícita*, al menos en las versiones más usuales del análisis. En estas versiones, las relaciones causales entre hechos particulares se dan *en virtud de que los hechos ejemplifican ciertas propiedades y de que hay una relación nómica entre esas propiedades*, esto es, ciertas leyes que las conectan.

11. Si no se acepta causalidad indeterminista, al defensor del análisis causalista sólo le queda, o bien rechazar que los sucesos indeterministas tengan explicación, o bien que haya genuinos sucesos indeterministas (por ejemplo, defendiendo que todo indeterminismo es sólo epistémico y que, metafísicamente hablando, todo suceso es determinista). Así, quien acepte que hay sucesos metafísicamente indeterministas,

de explicaciones aceptadas en ciencia cuya naturaleza causal es discutible. Se aducen a veces los casos de la psicología, la sociología o la historia, pero su supuesto (por los críticos del causalismo) carácter no causal es (independientemente de la cuestión del indeterminismo) muy discutible. Más problemático para el causalista es el caso de la *mecánica relativista*, cuyo *geometrismo* es para muchos muy difícilmente interpretable causalmente y, en cualquier caso, usa principios explicativos sin interpretación causal clara, como el principio de equivalencia.

En cuanto a la generalidad y especificidad, el análisis causalista determina qué es lo general de los diversos tipos de explicación y qué es lo específico. Toda explicación es causal. La explicación de hechos particulares (deterministas, y también de los indeterministas si se resuelven los problemas que acabamos de señalar) es causal porque consiste en proporcionar información causal contextualmente relevante de la ocurrencia del hecho particular. Y la explicación de hechos generales (leyes) es causal en sentido derivativo: si el hecho general es no causal, meramente descriptivo (por ejemplo, las leyes de Kepler), su explicación consiste en derivarlo de otra regularidad que sí sea causal (por ejemplo, las leyes de la mecánica clásica); si el Exm es una regularidad causal (por ejemplo, la ley de los gases), su explicación consiste en derivarla de otra causal (por ejemplo, las leyes de la mecánica estadística) más relaciones de constitución entre ambas (por ejemplo, que un gas es un conjunto de moléculas).

Para muchos, la principal dificultad del análisis causalista tiene que ver con el último *desideratum*, esto es, el de elucidación filosófica. Para los críticos, el análisis causalista elucidada la noción de *explicación* a partir de otra noción, la de *causa*, tanto o más necesitada de elucidación que la anterior, esto es, “elucida lo oscuro mediante lo más oscuro”. Los causalistas pueden responder, o bien aceptando que su análisis parte del concepto de causa como primitivo, pero rechazando que eso sea filosóficamente pernicioso, o bien dando a su vez un análisis del concepto de causa. Entre estos últimos, algunos lo hacen en términos de otras nociones modales, modalidad que los críticos consideran tan problemática como el concepto de causa. Otros filósofos de orientación más empirista que consideran las nociones modales primitivas sospechosas, intentan otras estrategias en términos no modales. Ambas estrategias se enfrentan a sus propias dificultades¹².

que estos sucesos tienen explicación y que no hay causalidad indeterminista, debe abandonar la idea de que toda explicación es causal. Tal es el caso, por ejemplo, de quienes defienden que la mecánica cuántica es completa, explicativa y que no hay causalidad indeterminista.

12. Para una defensa de la primera, *cf.*, por ejemplo, Lewis (1973), que utiliza la noción de *dependencia contrafáctica*; y de la segunda, *cf.* Salmon (1981, 1984, 1994, 1997) y Dowe (1992, 1995), que utilizan la noción de *transferencia de magnitud conservativa*.

3.3 *Explicación y unificación*

La intuición básica que inspira el modelo de explicación como unificación es que explicar consiste en “reducir la cantidad de fenómenos independientes que tenemos que aceptar como últimos” (Friedman, 1974, p. 18). Aumentamos nuestra comprensión del mundo reduciendo el número de supuestos básicos que dan cuenta de los fenómenos. Las leyes de Newton explican las de Kepler porque, además de implicarlas, reducen la cantidad de regularidades que se aceptan independientemente unas de otras: antes de la explicación, las leyes de Kepler y, por ejemplo, la de Galileo, eran aceptadas independientemente unas de otras, después no; la reducción de la temperatura a la energía cinética media amplía todavía más ese proceso de unificación explicativa. Contrariamente, la mera conjunción de, por ejemplo, las leyes de Kepler con la de Boyle no es una explicación de las primeras porque no produce ese efecto unificador, no permite simplificar la cantidad de supuestos primitivos. Esta noción de explicación está esencialmente relativizada a un cuerpo *K* de creencias aceptadas en un momento dado, y exige una elucidación precisa de la *independiente aceptabilidad* entre creencias. Quien ha desarrollado con más detalle este modelo ha sido P. Kitcher mediante sus nociones de *patrón argumentativo* y “*almacén*” *explicativo* (*explanatory store*) (cf. Kitcher, 1981, 1989, 1993), caracterizando el poder unificador en función de: (i) la cantidad de fenómenos derivados por el conjunto de patrones argumentativos, (ii) el rigor de los patrones y (iii) el número de patrones.

El modelo unificacionista puede ser visto como una sofisticación del modelo inferencialista de Hempel. En el modelo de Hempel una explicación es una inferencia que cumple ciertas condiciones, condiciones que son básicamente *locales* (que Exs contenga leyes, etc.). Ello es insatisfactorio por presentar contraejemplos, básicamente inferencias que cumplen esas condiciones, pero de las que no aceptamos que sean explicaciones satisfactorias. El modelo unificacionista añade condiciones nuevas, pero de carácter *global*, esto es, condiciones que no afectan sólo a la inferencia en cuestión sino a su relación con otras inferencias, básicamente: que la inferencia permita simplificar el conjunto total de inferencias del sistema sin perder en cantidad de hechos explicados. Y así resuelve los contraejemplos a la suficiencia que se le plantean a Hempel. La estrategia general es mostrar que, dadas dos inferencias alternativas, será explicativa la que pertenezca a la sistematización más unificadora, esto es, más simple (sin perder en cantidad de hechos explicados), y que esta comparación arroja en (la mayoría de) los casos en consideración los resultados intuitivamente esperados. Por ejemplo, en los casos de simetría, como el del mástil, la altura del mástil explica la longitud de la sombra y no al revés, pues dadas dos sistematizaciones, una que contenga inferencias que parten de la altura

del mástil y conducen hasta la longitud de la sombra, y otra que contenga inferencias que proceden al revés, la primera es más unificadora que la segunda. Si la segunda no tiene otro tipo de inferencias, pierde algunas conclusiones pues no podrá establecer, por ejemplo, la altura de mástiles de noche, o en días nubosos, etc. Para recuperar esas conclusiones debería introducir nuevos patrones argumentativos. Análogamente se abordan los casos de efectos de causa común, irrelevancia y precedencia temporal (aunque en este último caso hay dificultades específicas).

Parece pues que, al menos en principio, el modelo unificacionista satisface relativamente bien los *desiderata* de adecuación extensional y de generalidad/especificidad. Sus principales retos tienen que ver con el *desideratum* de elucidación filosófica. En principio, parece que mantiene el beneficio de Hempel de analizar una noción oscura, la de *explicación*, mediante otras más claras (patrón argumentativo, simplicidad, etc.). Pero los detalles del análisis ponen de manifiesto varias dificultades o retos, principalmente dos. Por un lado, el análisis debería proporcionar un criterio adecuado de ponderación entre los parámetros (a)-(c) que determinan la simplicidad comparativa, lo que es extremadamente difícil incluso si hubiese criterios no problemáticos para cada uno de ellos por separado. Por mencionar sólo la simplicidad y la fuerza/alcance, es inmediato que estos dos criterios se contraponen: un modo de obtener un sistema muy simple, es decir, con muy pocos patrones, es rebajar la fuerza, por ejemplo, perdiendo fenómenos explicados, y viceversa; cuantos más fenómenos explicados, menos simple es el sistema. Ponderar ambos parámetros requiere un criterio de comparación entre ambos y no es conceptualmente fácil indicar en qué consistiría. El segundo reto es el de recoger un sentido suficientemente fuerte de la objetividad de las explicaciones, y hacerlo sin recursos conceptuales que arruinen el inicial beneficio elucidatorio del análisis¹³.

3.4 Causalidad versus unificación

Estos dos modos de ver la explicación *no son incompatibles* entre sí; cada uno ofrece un modo razonable de analizar la explicación. De hecho, pueden ser to-

-
13. La principal dificultad para ello proviene del hecho de que la simplicidad depende del vocabulario que aceptemos como primitivo. Podemos simplificar mucho el número de principios básicos si aceptamos como primitivos predicados “raros”, como “verdul” (*grue*). Para eliminar esta indeterminación, una posibilidad es exigir que los predicados primitivos sean *proyectables* o que denoten clases naturales. Pero ambas alternativas, u otras semejantes, hacen depender el análisis unificacionista de nociones como *proyectabilidad* o *clase natural*, tan necesitadas o más de la elucidación que la de explicación, con lo que se arruina el inicial beneficio elucidatorio que motivaba la línea hempeliana.

mados como representando dos aspectos diferentes pero compatibles de la explicación científica (Salmon, 1989, p. 183).

Los modelos causalista y unificacionista de la explicación científica parecen en principio contrapuestos, esto es, propuestas diferentes y enfrentadas de análisis de la noción de explicación científica. Sin embargo, la cita de Salmon sugiere que esta contraposición puede ser sólo aparente y que en realidad se trata de análisis compatibles (*cf.* también, por ejemplo, Salmon, 1992, p. 39 y 2002a p. 105). Hemos realizado en otro lugar un análisis exhaustivo de las diversas posibilidades de compatibilidad; para nuestros actuales fines será suficiente tener presentes las principales conclusiones que dicho análisis arroja (*cf.* Díez, 2002):

- No es conceptualmente iluminador sostener la complementariedad de dos nociones sobre la base exclusiva de relaciones extensionales *no acompañada de algún tipo de conexión intensional/conceptual*. Si ese es todo el sentido en que los análisis causalista y unificacionista son compatibles, entonces no es un sentido filosóficamente interesante. Permanecería en la oscuridad filosófica cómo es que ambos tipos de explicaciones son *explicaciones*.
- (i) La existencia de un núcleo conceptual común es, pues, la opción mínima aceptable. Que haya otra opción aceptable más fuerte depende de la viabilidad de la reducción conceptual de una a otra. Si ninguna de estas reducciones fuera viable, deberíamos contentarnos con esta compatibilidad mínima. En este caso, debería proporcionarse un análisis satisfactorio de ese núcleo conceptual común, y hacerlo en términos que no presupusieran ni el causalismo ni el unificacionismo. De momento no se ha proporcionado un análisis semejante.
- (ii) Es poco plausible que toda explicación científica sea *conceptualmente* explicación causal. Hay casos claros de explicación científica cuya naturaleza causal es, cuando menos, controvertida. Si no en historia, sociología o psicología, sí al menos en algunas partes de la física, como la mecánica cuántica o, más claramente, en la mecánica relativista, cuyas explicaciones geométricas o las que apelan al principio de equivalencia son plenamente aceptadas, independientemente de que su carácter causal sea, cuando menos, controvertido. Si ello es así, su aceptación como explicaciones no puede estar dependiendo conceptualmente de su carácter causal.
- (iii) No hay objeciones fuertes contra la idea de que toda explicación científica es, conceptualmente, unificatoria, aunque falta todavía precisar el sentido de “unificación” que haga a esta idea mínimamente precisa y a la vez plau-

sible. Si ello no fuera posible, deberíamos limitarnos a buscar un núcleo conceptual básico común.

La propuesta de análisis de la explicación científica como subsunción teórica ampliativa que sigue a continuación puede interpretarse como un intento de hacer preciso *un* sentido plausible de explicación unificatoria. Si su pretendida proximidad con la idea básica unificacionista resulta demasiado débil como para considerarse una propuesta en esa línea, entonces deberá tomarse como un intento de elucidar el núcleo conceptual común a que se limita la opción (iii).

4. Explicación científica como subsunción modelo-teórica ampliativa

Para presentar de modo adecuado las ideas centrales del análisis subsuntivista de la explicación científica, es preciso ofrecer antes brevemente un esbozo del análisis de las entidades a que tales explicaciones están vinculadas, a saber: las teorías científicas. Según el análisis subsuntivista, la explicación científica va asociada típicamente a las teorías científicas que representan cierto dominio de datos mediante modelos teóricos ampliados. En qué sentido la subsunción de los modelos de datos en modelos teóricos ampliados es explicativa, y en qué medida ese sentido de explicación es próxima al análisis unificacionista, es algo que depende de la estructura fina de tales subsunciones y, por tanto, de la estructura de las teorías en cuyo marco se llevan a cabo.

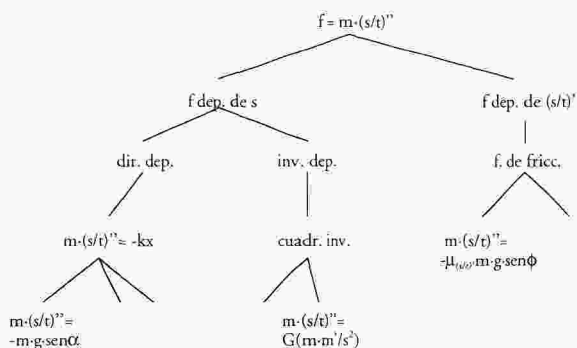
4.1 Las teorías como entidades modelo-teóricas reticulares¹⁴

- (i) *Análisis modelo-teórico.* Las teorías no se identifican mediante conjuntos de enunciados sino mediante conjuntos de modelos (aunque esos modelos son definidos mediante enunciados). Así, por ejemplo, para la identificación de la mecánica clásica (MC) serán relevantes los modelos del tipo $\langle P, t, s, m, f \rangle$ (P es un conjunto de partículas, por ejemplo, la tierra y la luna, t un intervalo temporal, s la función posición para las partículas en ese intervalo, m la función masa y f la función fuerza en ese intervalo) y otros definidos a partir de ellos.
- (ii) *Distinción T-teórico / T-no teórico.* Los términos/conceptos de una teoría T son los términos (primitivos) que aparecen en la formulación de las leyes de T . Un término es T -no teórico si su denotación en el modelo *se puede* determinar (medir) sin usar las leyes de T , y es T -teórico en caso contrario, si *todos* sus

14. Para una presentación un poco más detallada, cf. Díez y Moulines (1997), cap. 10 sec.5; para una presentación completa, cf. Balzer, Moulines y Sneed (1987).

- procedimientos de determinación presuponen alguna ley de T . En nuestro ejemplo, *masa* y *fuerza* son MC-teóricos, el resto MC-no teóricos.
- (iii) *Fenómenos, base empírica*. Los *modelos parciales* son los recortes T -no teóricos de los modelos de T , por ejemplo, en MC son del tipo $\langle P, t, s \rangle$. Éstos son los modelos que representan los fenómenos particulares que explica la teoría, su base empírica o *modelos de datos*. El conjunto de sistemas físicos *concretos* (por ejemplo, un niño columpiándose, la luna girando en torno a la tierra, ...) a los que se pretende aplicar la teoría es el *dominio de aplicaciones pretendidas I*.
- (iv) *Para "dar cuenta" de I, la teoría impone diversos tipos de constricciones formales a los modelos*. Las principales son las leyes. Las leyes que usan sólo términos de T definen un conjunto M de *modelos actuales*, por ejemplo, en MC las estructuras de tipo $\langle P, t, s, m, f \rangle$ que satisfacen además tales y cuales leyes. A veces algunas leyes usan conceptos de T y de otras teorías, son leyes "puente". Éstas también determinan un cierto conjunto de modelos en T , el conjunto L de *vínculos*. El último tipo principal de constricciones se debe a requisitos de coherencia que T impone, por ejemplo, que cuando una partícula aparezca en dos modelos m le asigne la misma masa en ambos. El conjunto que expresa esta restricción sobre los modelos es el conjunto C de *ligaduras*. Así, podemos expresar el conjunto total de constricciones teóricas sobre los modelos mediante la terna $K = \langle M, C, L \rangle$.
- (v) *Aserción empírica*. Las constricciones formales de K tienen efectos a nivel T -no teórico. Si "juntamos" las de M , C y L , obtenemos el *contenido teórico* CnT , y si de los modelos de CnT recortamos su parte T -teórica entonces obtenemos el *contenido empírico* de T , CnE . La aserción empírica de la teoría es que los sistemas físicos de I cumplen las constricciones de K , esto es, que los sistemas de I están entre los de CnE .
- (vi) *Redes teóricas*. No todas las constricciones formales están al mismo nivel, no todas son igualmente centrales o básicas para T . En MC, por ejemplo, la explicación de cualquier fenómeno usa la segunda ley de Newton, pero no toda explicación usa la ley de Hooke (para un muelle sí, para un planeta no), o la de gravitación (para un planeta sí, para un proyectil también, para un muelle no), o la del rozamiento, etc. Las constricciones vienen estructuradas por estratos, conformando una *red teórica*, empezando con las leyes (conjuntos de modelos) más generales y añadiendo después leyes más específicas en diversas direcciones para dar cuenta de fenómenos específicos (por ejemplo, en MC, una rama para fuerzas dependientes de la distancia, otra para fuerzas dependientes de la velocidad, etc.; una sub-rama de la primera

para fuerzas dependientes directamente de la distancia, otra para las dependientes inversamente, etc.). El siguiente gráfico representa esta estructura (los nodos inferiores son las *especializaciones* de sus superiores en la red):



En este gráfico, cada punto de la red tiene sus propias aplicaciones pretendidas, siendo las de los nodos superiores la reunión de las de sus correspondientes nodos inferiores/especializaciones. Con estas ideas podemos ya presentar lo fundamental del análisis de las explicaciones científicas como subsunciones de modelos de datos en modelos teóricos ampliados.

4.2 Explicación y subsunción teórica ampliativa

La noción central que recoge la idea de explicación dentro de este marco modelo-teórico es la de *subsunción* (*embedding*). Esta noción está presente ya en las diversas concepciones semánticas, y es central en el programa estructuralista desde sus mismos orígenes, ocupa un lugar preponderante en *An Architectonic* (Balzer, Moulines y Sneed, 1987) y ha sido desarrollada como propuesta de análisis de la explicación científica cercana al modelo unificacionista, principalmente por Bartelborth (1996, 1999, 2001). Vamos a presentar aquí la propuesta básicamente en la forma en que aparece en Bartelborth, aunque resaltando más que él el aspecto *ampliativo* de la subsunción para constituir una explicación.

La idea básica que se encuentra tras la noción de explicación como subsunción teórica es que explicar un fenómeno consiste en subsumir el fenómeno en un *patrón nómico*, esto es, en los términos modelo-teóricos vistos, en subsumir el fenómeno en una rama de la red teórica de constricciones nómicas K . Explicamos, por ejemplo, el movimiento de la luna (L) en torno a la tierra (T) durante cierto intervalo t , subsumiendo el modelo parcial $y = \langle \{T, L\}, s, t \rangle$ en las constricciones teóri-

cas CnT de cierta rama que comienza con la segunda ley de Newton, continúa con el principio de acción y reacción, y prosigue hasta la ley de gravitación (más *links* y ligaduras). Aquí “subsumir” significa mostrar que en CnT hay modelos $\langle \{T, L\}, s, t, m, f \rangle$ que cumplen tales restricciones y que tienen a y como parte MC-no teórica. Así, explicamos ese fenómeno cinemático mostrando que, si la tierra y la luna tuviesen tales masas y estuvieran sometidas a tales fuerzas que se comportan con m , s y t del modo específico que tales constricciones determinan, entonces su movimiento debería ser el que de hecho es. Y análogamente con la explicación del movimiento de un objeto en el extremo de un muelle, o de un niño en un columpio, o de un cuerpo descendiendo en un plano inclinado con rozamiento, etc.

Es en este sentido, por tanto, en el que explicamos los fenómenos empíricos contenidos en el dominio I de aplicaciones propuestas subsumiéndolos en el núcleo nómico K de constricciones de la correspondiente rama de la red teórica. Nótese que la explicación de los fenómenos T -no teóricamente identificados es proporcionada por los constituyentes T -teóricos de los modelos. Es *porque* en el sistema están “presentes” esos constituyentes T -teóricos comportándose con los T -no teóricos como las constricciones establecen, por lo que entendemos el comportamiento de la parte T -no teórica del sistema. Para que haya explicación por subsunción, por tanto, tiene que haber *extensión* o *ampliación teórica*. Aunque usualmente es lo que ocurre, no toda teoría genera extensión teórica, y aquellas que no la generen no podrán considerarse *explicativas*. Por ejemplo, la cinemática galileana, o la astronomía kepleriana¹⁵, no lo son. Estas son teorías puramente descriptivas, en las que las “leyes” describen simplemente ciertas regularidades empíricas sin añadir la teoría aparato T -teórico nuevo que refiera a entidades nue-

15. La astronomía kepleriana entendida meramente como (los modelos de) las tres leyes estrictamente cinemáticas que (con independencia ahora de la mayor o menor fidelidad histórica) se han popularizado como “las leyes de Kepler”, a saber: (1) los planetas siguen órbitas elípticas con el sol ocupando uno de los focos; (2) “barren” áreas iguales en tiempos iguales, y (3) el cuadrado del período de cada planeta es proporcional al cubo de su distancia media al sol). Si “aumentamos” esta teoría con algo como “la constante de Kepler” (que corresponde aproximadamente al cociente entre la masa del planeta y la del sol), sí tenemos una teoría explicativa. O si incluimos la hipótesis neopitagórica del *anima motrix*. Otra cuestión es si el *explanandum* no es el movimiento real de los planetas sino su movimiento aparente a través de la eclíptica, por ejemplo, las trayectorias retrogradantes de ciertos puntos en el cielo visible. En tal caso, esas tres leyes sí se pueden considerar explicativas, pero no obviamente “del movimiento *real* de los planetas”.

vas diferentes a las involucradas en los modelos de datos. En estas teorías los modelos parciales y los modelos posibles coinciden, pues no hay componentes *T*-teóricos que “recortar”. Eso no quiere decir que los fenómenos descritos por esas teorías no puedan ser explicados. Lo que sucede es que en estos casos el proceso de explicación es más complejo y “derivado” del anterior a través de relaciones inter-teóricas reductivas que mantienen con otras teorías, por ejemplo, en estos dos casos con la mecánica clásica.

Éste es, pues, el esquema básico del análisis de la explicación científica como subsunción teórica ampliativa. Explicar un fenómeno particular consiste en subsumirlo en un modelo teórico *ampliado* en el marco de una red teórica. ¿Qué tal le va a este modelo respecto de los tres criterios de adecuación que establecimos para todo análisis de la explicación científica? Por lo que se refiere a la elucidación filosófica, este análisis no utiliza otros conceptos tan o más oscuros que el de explicación. El análisis depende, básicamente, de los conceptos de *subsunción ampliativa* y de *red teórica*, conceptos que elucidan a su vez mediante los de *T-teoricidad*, *modelos de datos*, *aplicación pretendida*, *especialización*, y otros que elucidan a su vez de modo claro. Parece pues que este análisis no elucidan “lo oscuro mediante lo más oscuro”.

En cuanto a la especificidad/generalidad, parece que también lo cumple satisfactoriamente. El modelo que hemos dado se aplica esencialmente a la explicación de hechos particulares, independientemente de que éstos sean deterministas o indeterministas. Las explicaciones de hechos particulares deterministas se diferencian de los indeterministas simplemente en que las teorías/modelos en que se subsumen son deterministas en el primer caso y probabilistas en el segundo. El tipo de subsunción es el mismo; el tipo de modelo en que se subsume es diferente.

¿Cómo da cuenta este modelo de las explicaciones de *hechos generales*? En este fenómeno siempre está involucrado algún tipo de reducción inter-teórica, y la reducción inter-teórica se elucidan, indirectamente, en términos de subsunción. Ahora bien, el tipo de reducción no es siempre el mismo. Hay que distinguir por lo general dos tipos de casos, según los hechos generales explicados sean a su vez explicativos o no lo sean. Las leyes galileanas del movimiento, o las leyes de Kepler son de este segundo tipo. Las leyes de la mecánica del sólido rígido o las de la termodinámica son del primero. La reducción de la cinemática galileana y la astronomía kepleriana a la MC es una reducción exclusivamente a nivel empírico o aplicativo. Esta reducción consiste simplemente en que la teoría reductora incluye entre su dominio de aplicaciones pretendidas *I* modelos parciales que expresan las regularidades no explicativas de las teorías reducidas, modelos que son explicados mediante subsunción ampliativa en la teoría

reductora. En este caso, pues, se trata de reducción explicativa pero sin *sustitución explicativa*. Un fenómeno *general* no explicado se explica en el marco de la reductora. Por otro lado, la reducción de la mecánica del sólido rígido o la termodinámica a la MC es una reducción teórica o de *sustitución explicativa*, pues estas teorías realizaban sus propias explicaciones por subsunción ampliativa de determinados fenómenos. La reducción explicativa en este caso consiste básicamente en correlacionar los modelos parciales y los completos de ambas teorías de modo tal que, el que ciertas aplicaciones de la reductora sean subsumibles en ciertas constricciones K implica que las correspondientes aplicaciones de la reducida sean subsumibles en las correspondientes constricciones K' . Esto es, una subsunción-explicación sustituye a otra. Y no sólo la sustituye, sino que, en cierto sentido, la explica. El que ciertas subsunciones-explicaciones (de la reductora) funcionen, explica que ciertas otras subsunciones-explicaciones (de la reducida) funcionen. Esta reducción entre explicaciones no es literalmente del mismo tipo que las subsunciones, pero se analiza en términos de subsunciones.

Por último, por lo que se refiere a la adecuación extensional, nos vamos a restringir de momento a las explicaciones *científicas* (o a la eventual versión científica de las explicaciones de los ejemplos); cómo se enfrenta este modelo con las explicaciones precientíficas ordinarias se discutirá más adelante. Así, la bajada de la presión explica la tormenta pero no la explica la bajada del barómetro porque en la subsunción teórica del modelo de datos que representa la tormenta no desempeña ningún papel la bajada del barómetro pero sí el descenso de la presión. Y análogamente para los casos de la altura del mástil y del no embarazo de Juan. Quizás se objete que siempre es posible construir artificialmente tales subsunciones y que la cuestión es cómo excluirlas. La respuesta en este caso es sustancialmente la misma que la unificacionista: esas subsunciones serían insatisfactorias, o en cualquier caso más insatisfactorias que las otras, atendiendo a criterios de unificación y simplicidad. Un caso diferente es el del eclipse en cierto momento t , y en general el de la asimetría temporal, por ejemplo, en mecánica y otras teorías “reversibles”. En MC son igualmente subsumibles ampliativamente tanto el modelo que representa la trayectoria del sistema tierra/sol/luna entre t' y t ($t' < t$), como el modelo que representa la trayectoria del sistema tierra/sol/luna entre t y t'' ($t < t''$). Ambas subsunciones nos sirven para poder “predecir” el eclipse en el primer caso y “retrodecirlo” en el segundo. Pero recordemos que ahora explicación no es mera derivación. En los dos casos hay subsunción, y lo que se subsume en ambos casos es cierta trayectoria cinemática, en el primero, una que acaba en el eclipse y en el segundo, otra que empieza en él. *En tanto que trayectorias de un sistema durante un período de tiempo*, ambas son igualmente subsumibles y por tanto igualmente ex-

plicadas. Y ello coincide perfectamente con nuestras intuiciones. Otra cuestión es cuál de las dos subsunciones, o si las dos, pueden considerarse explicaciones del eclipse. Eso es algo que el modelo parece dejar de momento indeterminado¹⁶.

En cuanto a la adecuación extensional, otra cuestión diferente es si cabe esperar que todas las explicaciones científicas se adecúen al patrón subsuntivista, incluidas las explicaciones en algunas ciencias humanas, particularmente en la historia. Ésta es una cuestión que entronca con la tradicional polémica entre explicación y comprensión, pero en la medida en que sea legítimo hablar genuinamente de explicaciones en las ciencias historiográficas, no veo motivos de principio por los que no se puedan adecuar a este esquema. Un problema independiente es si las explicaciones en historia son siempre tan complejas estructuralmente como este modelo presupone, pues ello tiene que ver con la adecuación de este modelo a las explicaciones pre o protocientíficas, que hemos dejado para las consideraciones finales.

4.3 Subsunción, unificación y causalidad

El modelo subsuntivista presenta ciertos rasgos que lo aproximan a la tradición inferencialista, y en especial a la versión unificacionista del inferencialismo. T. Bartelborth ha hecho énfasis en los aspectos del subsuntivismo que más manifiestan un “espíritu” unificacionista. Entre ellos se destacan los siguientes:

- (a) Los fenómenos no se explican “suelos” sino “en grupos”. Los elementos de *I* no son modelos parciales suelos sino grupos de ellos; por ejemplo, la Luna dando vueltas a la Tierra, junto con Venus dando vueltas al Sol, junto con..., o un niño en un columpio, junto con el péndulo de un reloj, junto con..., etc.

16. Ante ellas caben dos alternativas. La primera, reforzar el modelo exigiendo que, cuando un modelo de datos que representa un proceso se subsume en un modelo teórico, se puede decir derivativamente que los eventos anteriores del proceso explican los posteriores, pero no a la inversa. Ello supondría introducir primitivamente la precedencia temporal en este sentido derivativo de explicación. La segunda opción es dejar genuinamente indeterminada la asimetría como una asimetría en la explicación y dar cuenta de nuestras intuiciones en otros términos. Por ejemplo, la razón de que consideremos que las posiciones en t' explican (derivativamente) las posiciones en (t) , pero que las de (t') no, podría ser que confundimos o “mezclamos” dos cuestiones distintas, una la cuestión genuina de la explicación, en la que no habría diferencias, y otra la cuestión de la *predicción*, en la que sí hay diferencia y ésta tiene que ver obviamente con el orden temporal. Éste es un aspecto de momento abierto que el modelo subsuntivista debe elaborar.

- (b) La explicación es tanto mejor cuanto mejor es el balance entre *aplicación* e *información*. La aplicación es tanto mayor cuanto mayor es la parte de I exitosa. La información es tanto mayor cuanto más restringe K , esto es, cuanto menor es CnT . Y estos dos parámetros compiten entre sí: se puede ampliar mucho la aplicación disminuyendo mucho la información (cuanto menos constricciones imponga K , más fácil es que las aplicaciones de I estén efectivamente en CnE), y viceversa, hasta los casos extremos de aplicación total por información nula (tautológica) o información máxima con aplicación nula (Bartelborth, 2001, pp. 9-11).
- (c) La explicación es *integradora* u *orgánica*: (i) La explicación de unos fenómenos en una rama de la red está integrada con la de otros en otras ramas a través de la dependencia de ambas de las partes comunes de sus respectivas ramas, al menos las constricciones generales superiores (por ejemplo, la segunda ley de Newton) comunes a toda la red. (ii) La explicación conecta modelos entre sí a través de las condiciones de ligadura C . (iii) La explicación conecta con otras teorías a través de los vínculos inter-teóricos L que forman parte de las restricciones de K .
- (d) Dentro de una misma teoría, se produce progreso explicativo cuando fenómenos que requieran explicaciones diferentes se explican ahora del mismo modo. En términos de redes teóricas, cuando la red se simplifica *sin perder aplicatividad*. Para que ello represente un progreso efectivo se deben excluir casos en los que simplemente se “conjuntan” ramas diferentes. Para evitar esa situación, la red ha de ser tal que en ningún extremo terminal $\langle K, I \rangle$ de una rama sea posible dividir K, I respectivamente en K_1, K_2, I_1, I_2 tales que I es subsumible bajo K y I_1 es subsumible bajo K_1 y I_2 es subsumible bajo K_2 (modificación de de Bartelborth, 2001, p. 10).

Si estos rasgos se consideran una precisión suficiente y plausible del núcleo de la intuición unificacionista, entonces el análisis subsuntivista podría proporcionar la versión de unificacionismo que apoyase la prioridad conceptual de unificacionismo frente a causalismo. Si no se consideran suficientes y son esenciales al unificacionismo intuiciones adicionales, entonces el subsuntivismo sustentaría tan sólo la alternativa más débil, esto es, proporcionaría un núcleo conceptual común a toda explicación científica, ya sea unificadora o causal. Para ello es preciso que el subsuntivismo sea compatible con la existencia de explicaciones causales. Y efectivamente lo es: una explicación/subsunción es causal cuando las entidades T -teóricas con las que T amplía los modelos de datos dando lugar a modelos teóricos, son entidades con *poder causal* (sea esto lo que sea).

Así, no nos preguntamos si hay operando mecanismos causales para poder decir que explicamos un determinado fenómeno. Quizás los haya a veces, incluso pudiera haberlos siempre, pero en este marco eso no es conceptualmente relevante para hablar de explicación. Se explica un fenómeno cuando se subsume ampliativamente en la correspondiente rama de la red. La explicación será *además* causal si alguno de los constituyentes *T*-teóricos con los que ampliamos los modelos *T*-no teóricos tiene *potencia causal*, o alternativamente, si las constricciones-leyes de *K* son leyes causales. Muchas veces así es, como en la explicación mecánico-clásica de los fenómenos cinemáticos. Pero en algunos casos puede no ser así, o es cuando menos controvertido y ello no impide hablar de explicación. Algunos, por ejemplo, defienden que sólo las leyes de sucesión son causales, y las de coexistencia no. Desde la perspectiva subsuntivista, cualquiera de ellas son explicativas mientras intervengan en subsunciones ampliativas, independientemente de su eventual carácter causal. Quizás la explicación del movimiento del globo en un avión apelando al principio de equivalencia no sea causal, o la de la conducta en términos de deseos y creencias, o las mecánico-cuánticas, pero ello no obsta para que las calificemos de explicaciones genuinas sin esperar resolver la controversia, y lo hacemos porque independientemente de su estatuto causal, también son casos de subsunción teórica ampliativa

5. Consideraciones finales

La noción de *explicación científica* como *subsunción teórica ampliativa* expresa un sentido claro de “explicación”, acorde con los casos paradigmáticos de explicación científica, y que se adecua satisfactoriamente a los *desiderata* con que se ha de afrontar todo análisis de la explicación, o al menos más satisfactoriamente que los modelos anteriores. Respecto de éstos, el análisis subsuntivista permite preservar y combinar sus virtudes sin cargar con sus deficiencias:

- (a) *Análisis inferencialista nómico*. El subsuntivismo comparte con él la intuición de que explicar supone algún tipo de “esperabilidad” o “encaje”, así como que involucra leyes. No se compromete, sin embargo, con que la explicación sea siempre inferencial en sentido estricto, y tampoco con que sea local.
- (b) *Análisis causalista*. El subsuntivismo comparte con él la intuición de que explicar supone siempre “poner algo ontológicamente nuevo sobre el tapete”. Pero discrepa en que ello tenga que ser siempre causal para poder hablar legítimamente de explicación. A veces, incluso la mayoría de las veces, lo es, pero no es conceptualmente necesario.
- (c) *Análisis unificacionista*. El subsuntivismo comparte con él la intuición de que la explicación es básicamente no-local, y que es tanto mejor cuanto más

orgánica/unificadora sea (sin perder alcance). Pero discrepa en las versiones estrictamente inferencialistas del unificacionismo (Kitcher), y está en cuestión si los rasgos unificacionistas que presenta son suficientes para asentar el unificacionismo como la alternativa conceptualmente más general.

El modelo unificacionista representa, en mi opinión, una propuesta prometedora de análisis de la explicación científica, que recoge lo mejor de las anteriores, evitando sus deficiencias. El mayor problema con que se enfrenta deriva de su eventual generalización a todo tipo de explicación, incluida la explicación no científica ordinaria. Por un lado, parece que nuestra noción de explicación es tal que hay un núcleo conceptual común compartido por las explicaciones científicas y las precientíficas. Por otro, esta noción de explicación científica como subsunción teórica ampliativa parece difícilmente extrapolable a las explicaciones precientíficas, que carecen de la complejidad teórica presupuesta por el subsuntivismo. Ante esta situación caben tres opciones:

Una opción es renunciar a la idea de que hay un núcleo conceptual común a la explicación científica y a la precientífica. Ello parece poco deseable, fundamentalmente porque los casos paradigmáticos asociados a la expresión “explicación” son precientíficos, y por tanto dejaríamos en el misterio por qué las explicaciones científicas son denominadas “explicaciones”.

Otra opción menos drástica es defender que hay un núcleo conceptual común a la explicación científica y a la precientífica, pero que dicho núcleo no lo da la noción de subsunción teórica ampliativa y está todavía por analizar/descubrir. En ese caso, habríamos logrado dar un análisis satisfactorio de la noción de *explicación científica*, pero dicho análisis no arrojaría luz alguna sobre ese núcleo conceptual común, y por tanto dejaríamos también en este caso en el misterio por qué las explicaciones científicas son denominadas “explicaciones”, al menos provisionalmente, hasta que no diéramos con dicho núcleo conceptual común.

La última opción es debilitar un poco las fuertes condiciones estructurales presupuestas en el análisis, de modo que las explicaciones precientíficas se adecuen a ellas, siquiera vaga o gradualmente, pero conserven todavía el núcleo de las intuiciones subsuntivistas. Tal como las hemos expuesto, el modelo subsuntivista combina dos ideas: (i) explicar científicamente es subsumir el *explanandum*/modelo de datos en un *explanans*/modelo teórico que incorpora material ontológico nuevo, y (ii) esta subsunción tiene lugar en el marco de una red jerarquizada y conectada de constricciones sobre la relación entre el material nuevo y el ya existente en los modelos de datos. De estas dos ideas, la más esencial al subsuntivismo parece ser la primera, y también a la que más parece que se pueden adecuar las explicaciones precientíficas. De este modo, podríamos rela-

jar nuestras condiciones sin exigir que la red de constricciones sea siempre tan compleja y bien estructurada como en una teoría científica desarrollada. Las explicaciones precientíficas serían entonces subsunciones teóricas en un sentido de “teórico” que no presupondría la complejidad de una red teórica, sino tan sólo la introducción del aparato conceptual nuevo respecto del utilizado en la descripción de los datos. Ello supondría cierto grado de vaguedad o variabilidad en la “bondad” de las explicaciones, siendo tanto mejores las explicaciones cuanto más se aproximaran al modelo complejo, lo cual parece en principio una consecuencia deseable. Aunque las explicaciones precientíficas fuesen paradigmáticamente más básicas, también serían más vagas e imperfectas y su optimización daría lugar a las explicaciones científicas. El único problema con esta opción es que para que la noción de subsunción no sea vacía hay que mantener alguna fuerza en las constricciones, y esa fuerza depende típicamente de la interconexión de las diferentes constricciones, de modo que si se relajan totalmente dichas conexiones la subsunción puede devenir trivial. La cuestión es cuánto se pueden relajar sin trivializar la subsunción y si esas constricciones mínimas no triviales pueden satisfacerse plausiblemente en las explicaciones no científicas. Éste es el principal reto al que se enfrenta el subsuntivismo si pretende proporcionar una teoría general de la explicación.

Bibliografía

- Balzer, W., Moulines, C. U. y Sneed, J. D. (1987): *An Architectonic for Science. The Structuralist Program*, Dordrecht: Reidel.
- Bartelborth, T. (1996): “Scientific Explanation”, en: Balzer, W. y Moulines, C. U. (eds.), *Structuralist Theory of Science*, Berlín: Walter de Gruiter, pp. 23-43.
- Bartelborth, T. (1999): “Coherence and Explanation”, en: *Erkenntnis*, 50, pp. 209-224.
- Bartelborth, T. (2002): “Explanatory Unification”, en: *Synthese* 130-131, pp. 91-107.
- Díez, J. A. (1998): “Hacia una teoría general de la representación científica”, en: *Theoria*, 13/31, pp. 113-139.
- Díez, J. A. (2000): “Structuralist Analysis of Fundamental Measurement Theories”, en: Balzer, W., Sneed, J. D. y Moulines, C. U. (eds.), *Structuralist Knowledge Representations. Paradigmatic Examples*, Poznan Studies 75, Ámsterdam: Rodopi, pp. 19-49.
- Díez, J. A. (2002a): “Explicación, unificación y subsunción”, en: W. González (ed.), *Diversidad de la explicación científica*, Barcelona: Ariel, pp. 73-93.

- Díez, J. A. (2002b): "A Program for the Individuation of Scientific Concepts", en: *Synthese* 130-131, pp. 13-48.
- Díez, J. A. y Moulines, C. U. (1997): *Fundamentos de filosofía de la ciencia*, Barcelona: Ariel.
- Dowe, P. (1992): "Wesley Salmon's Process Theory of Causality and the Conserved Quantity Theory", en: *Philosophy of Science*, 59, pp. 195-216.
- Dowe, P. (1995): "Causality and Conserved Quantities: A Reply to Salmon", en: *Philosophy of Science*, 62, pp. 321-333.
- Friedman, M. (1974): "Explanation and Scientific Understanding", en: *The Journal of Philosophy*, 71, pp. 5-19.
- Hempel, C. G. (1965): *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science*, New York: Free P.
- Hempel, C. G. y Oppenheim, P. (1948): "Studies in the Logic of Explanation", en: *Philosophy of Science*, 15, pp. 567-579.
- Kitcher, P. (1976): "Explanation, Conjunction and Unification", en: *The Journal of Philosophy*, 73, pp. 207-212.
- Kitcher, P. (1981): "Explanatory Unification", en: *Philosophy of Science*, 48, pp. 507-531.
- Kitcher, P. (1989): "Explanatory Unification and the Causal Structure of the World", en: Kitcher, P. y Salmon, W. (eds.), *Scientific Explanation*, Minneapolis: U. Minnesota P., pp. 410-505.
- Kitcher, P. (1993): *The Advancement of Science*, Oxford: Oxford U.P.
- Kitcher, P. y Salmon, W. (1987): "Van Fraassen on Explanation", en: *The Journal of Philosophy*, 84, pp. 315-330.
- Kitcher, P. y Salmon, W. (eds.) (1989): *Scientific Explanation*, Minneapolis: U. Minnesota P. pp. 410-505.
- Lewis, D. (1973): "Causation", en: *The Journal of Philosophy*, 70, pp. 556-567.
- Lewis, D. (1986): "Causal Explanation", en: *Philosophical Papers II*, Oxford: Oxford U.P., pp. 214-240.
- Rubén, D. (1990): *Explaining Explanation*, New York: Routledge.
- Salmon, W. (1981): "Causality: Production and Propagation", en: Asquith, P. D. y Giere, R. (eds.), *PSA 1980*, Philosophy of Science Association, East Lansing MI, pp. 49-69.
- Salmon, W. (1984): *Scientific Explanation and the Causal Structure of the World*, Princeton NJ: Princeton U.P.
- Salmon, W. (1989): "Four Decades of Scientific Explanation", en: Kitcher, P. y Salmon, W. (eds.), *Scientific Explanation*, Minneapolis: U. Minnesota P., pp. 3-219.

- Salmon, W. (1992): "Scientific Explanation", en: Salmon, M. *et al.*, *Introduction to the Philosophy of Science*, Prentice Hall, Englewood Cliffs NJ, pp. 7-41.
- Salmon, W. (1994): "Causality Without Counterfactuals", en: *Philosophy of Science*, 61, pp. 297-312.
- Salmon, W. (1997): "Causality and Explanation: A Reply to Two Critiques", en: *Philosophy of Science*, 64, pp. 461-477.
- Salmon, W. (1998): *Causality and Explanation*, New York: Oxford U. P.
- Salmon, W. (2001a): "Causal *versus* Noncausal Explanation" (manuscrito).
- Salmon, W. (2001b): "The Structure of Causal Explanation" (manuscrito).
- Van Fraassen, B. (1977): "The Pragmatics of Explanation", en: *American Philosophical Quarterly*, 14, pp. 1143-1150.
- Van Fraassen, B. (1980): *The Scientific Image*, Oxford: Clarendon P.

Límites de la explicación como subsunción

Profesor del Departamento de Filosofía
de la Universidad Nacional de Colombia.

1. Explicación como subsunción

En su ponencia “La explicación científica: causalidad, unificación y subsunción teórica”, el profesor Díez sostiene que podemos tener un sentido claro de la noción de “explicación” entendida básicamente como *subsunción*, es decir, como integración de un fenómeno en una red teórica (en un patrón nómico). Esta propuesta puede ser vista como una forma de hacer eco a la tesis compatibilista de W. Salmon según la cual, el modelo causal de la explicación y el modelo unificacionista no constituyen realmente una disyuntiva, sino que deberían ser vistos como aspectos compatibles y complementarios a la hora de dar cuenta del concepto de explicación. En la perspectiva del profesor Díez, la compatibilidad de los modelos se da en términos de una dependencia conceptual del modelo causal con relación al unificacionista, según la cual toda explicación causal presupone una explicación como unificación. En este sentido, en el compatibilismo propuesto se entiende que el modelo de explicación científica como subsunción está más cerca del modelo unificacionista que del modelo causal. En el modelo unificacionista la explicación se entiende como reducción del número de supuestos básicos que dan cuenta del fenómeno. Por su parte, en el modelo de explicación como subsunción, la explicación se entiende como una especie de reducción del fenómeno (o de otras explicaciones) a una rama de la red teórica que está gobernada por ciertas constricciones formales. En ambos casos se trata, pues, de dar cuenta del fenómeno a partir de unos presupuestos teóricos en el contexto de los cuales el fenómeno es considerado como un ejemplar (*an instance*) de las regularidades que establece la teoría.

El modelo de explicación como subsunción también hace suyo el análisis de Van Fraassen acerca de los contextos explicativos, según el cual, la identificación de una explicación requiere que se determinen: (a) una pregunta tipo *por qué* o

“tema”; (b) una clase de contraste que expresa el aspecto por el que se pide explicación; y (c) una relación de relevancia explicativa que selecciona las razones que cuentan como explicativas en un determinado contexto. Adicionalmente, hay algunos presupuestos metateóricos, como la concepción “modelo-teórica” de la ciencia, según la cual, las teorías deben ser concebidas como conjuntos de modelos (más que como conjuntos de enunciados); la distinción teórico / no-teórico; la noción de base empírica en la ciencia; la idea de que las teorías establecen ciertas constricciones formales y, finalmente, la idea de que estas constricciones determinan redes teóricas. Una última idea básica en la noción de explicación como subsunción es la de *extensión* o *ampliación teórica*. De una teoría puede decirse que es explicativa, si permite una ampliación teórica, es decir, si hace posible que el comportamiento de un fenómeno (la parte no teórica del sistema en cuestión) se pueda comprender a la luz de los componentes teóricos del modelo y de las constricciones nómicas que éste establece. Se trata, pues, en estricto rigor, de un modelo de explicación como *subsunción teórica ampliativa*¹.

2. Explicación: causalidad, unificación y subsunción

Para tener algunos elementos de juicio acerca de la plausibilidad de la propuesta del modelo de explicación como subsunción teórica ampliativa, vale la pena considerar, aunque sea esquemáticamente, algunos elementos de la discusión entre causalismo y unificacionismo. Esto básicamente porque, como ya se ha señalado, el modelo propuesto se plantea como un modelo compatibilista con respecto a estas dos versiones de la explicación. En primer lugar, conviene anotar que la objeción de fondo al modelo causal proviene del contexto del empirismo y consiste en denunciar una infortunada mezcla de compromisos metafísicos con la solución de ciertos problemas epistemológicos. Así, desde la perspectiva de esta crítica, resulta imposible tratar de arrojar luces sobre el concepto de explicación si las bases del intento se sumergen más allá de lo empíricamente contrastable. Lo que pretenden denunciar los críticos empiristas del modelo causal es que un concepto epistemológico como el de *explicación* parece depender crucialmente de un concepto metafísico como el de *causalidad*. Para no sucumbir ante esta crítica, la estrategia de algunos teóricos consiste en redefinir el problema de la relación causal, convirtiéndolo en un análisis de las *relaciones entre los enunciados* que describen el fenómeno a explicar (el efecto) y los que describen las condiciones antecedentes específicas (las causas). Así, por ejemplo, en el modelo hempeliano de “cobertura de ley” (*the covering law model*) o modelo nomológico

1. Aunque en adelante me referiré a él como “modelo de explicación como subsunción”, no se debe perder de vista el aspecto ampliativo de la propuesta.

deductivo, la estrategia consiste en concebir la explicación como una relación entre oraciones; una de las cuales llamamos *explanandum* (que describe el fenómeno por explicar) y otras a las que llamamos *explanans* (que describen (a) las condiciones antecedentes específicas y (b) algunas leyes generales). Al situar al enunciado que describe el fenómeno por explicar como la conclusión de un argumento y a los enunciados que describen las condiciones antecedentes específicas como sus premisas, las relaciones causales quedan ahora puestas en el contexto de la deductibilidad. Con esta jugada, el modelo nomológico deductivo puede ser visto como una versión sofisticada del modelo causal (cf. Gaspar, 1991, p. 291). Un aspecto interesante en este punto es que el modelo nomológico deductivo es claramente un modelo de explicación como subsunción nómica. “El requisito decisivo para toda explicación sólida –dice Hempel– es que subsuma el *explanandum* en leyes generales” (Hempel, 1965 [1988, p. 259]). Quizá la formulación explícita de este requisito debería advertir que la intención de conciliar el causalismo con el unificacionismo tiene ya antecedentes muy importantes. En efecto, P. Kitcher en su “Explanatory Unification” (1991) advierte que en el empirismo contemporáneo, la explicación como unificación es una visión no oficial que está, sin embargo, presente en el modelo nomológico deductivo (cf. Kitcher, 1991, pp. 329-30). Así, el propósito que Kitcher persigue es el de explorar los beneficios de esta visión no oficial, para mostrar que es más promisoría que la visión estándar del modelo nomológico deductivo. Para Kitcher, la estrategia del modelo nomológico deductivo no es concluyente porque no puede dar cuenta de dos aspectos fundamentales de la explicación. En primer lugar, no puede dar cuenta de cómo es que la explicación científica logra ampliar nuestra comprensión del mundo. Ésta es una objeción mayor, pues pone al modelo nomológico deductivo en una encrucijada. Si explicar consiste en situar al *explanandum* como conclusión de un argumento deductivo, entonces, un fenómeno no queda explicado a menos que se garantice que el enunciado que lo describe se sigue de las premisas. Los argumentos deductivos garantizan esta relación, pero no son ampliadores de conocimiento, pues la conclusión está conceptualmente implícita en las premisas. Son los argumentos inductivos los que realmente permiten ampliar el conocimiento. Pero los argumentos inductivos no garantizan que la conclusión se siga de las premisas; lo que, en términos de la explicación, significa que las premisas usadas no constituyen la única explicación posible del *explanandum*. Así, cuando este modelo puede garantizar la explicación, no puede garantizar la ampliación del conocimiento; y cuando puede garantizar la ampliación del conocimiento, no puede garantizar la explicación. En segundo lugar, Kitcher afirma que el modelo nomológico deductivo tampoco puede dar cuenta de cuál es el papel que cumple el *poder explicativo* de las teorías a la hora de resolver las disputas entre teorías rivales. El *poder explicativo* es

entendido en este caso como la capacidad de explicar una cierta diversidad de fenómenos a partir de un mismo patrón argumentativo. Ésta resulta ser una virtud de las teorías que cumple una función muy importante a la hora de tener que escoger entre teorías rivales.

Sin embargo, la propuesta unificacionista de Kitcher no está exenta de dudas. En “The Pragmatics of Explanation” (1991), Bas van Fraassen señala algunas críticas a la idea del *poder explicativo* como una virtud o una propiedad que tengan las teorías en sí mismas. Desde el punto de vista de Van Fraassen, no hay una clara formulación de la relación que hay entre el poder explicativo de las teorías y el hecho de que éstas sean verdaderas. En esa medida, la noción de poder explicativo debería ser tratada sobre la base de criterios pragmáticos, es decir, sobre la base de la identificación de los intereses que orientan al científico o a la comunidad científica en un determinado momento. La respuesta de los unificacionistas a este respecto señala que algunos casos de la historia de la ciencia muestran que la adopción de ciertos programas de investigación, más que a criterios puramente pragmáticos, se ha debido a la superioridad del *poder explicativo* de tales programas.

Éstos son, pues, algunos elementos de la discusión entre causalismo y unificacionismo. Pero más que entrar en los detalles de este debate, lo que me interesa es tratar dos relaciones que quedan sugeridas en él. La primera es la relación entre *subsunción* y *ampliación teórica*. La segunda, es la relación entre *poder explicativo* y *selección teórica*. Se trata pues de explorar en este punto concreto los alcances del modelo propuesto de explicación como subsunción teórica ampliativa.

El problema de la relación entre “subsunción” y “posibilidad de ampliación teórica” corresponde al problema de explicar nuevos fenómenos con viejas teorías o, para formularlo en el lenguaje de Kuhn, al problema de cómo dar cuenta de las anomalías. Por otra parte, el problema acerca de si el *poder explicativo* de las teorías puede contar como evidencia para resolver las disputas entre teorías igualmente explicativas, puede ser formulado en términos de si la evaluación del poder explicativo de una teoría puede hacerse con criterios metateóricos, o si tal evaluación es, más bien, necesariamente intrateórica.

Para efectos de ordenar la exposición voy a partir de dos problemas que Van Fraassen considera básicos acerca de la explicación científica. El primero es cómo describirla, es decir, cómo establecer las condiciones de acuerdo con las cuales algo es explicado, problema que puede sintetizarse en la pregunta: ¿Qué es explicar? El segundo problema es el de establecer por qué creemos que las teorías deben ser explicativas, lo que se reduce a la pregunta: ¿Por qué es una virtud la explicación?

- Díez, J. A. (2002b): "A Program for the Individuation of Scientific Concepts", en: *Synthese* 130-131, pp. 13-48.
- Díez, J. A. y Moulines, C. U. (1997): *Fundamentos de filosofía de la ciencia*, Barcelona: Ariel.
- Dowe, P. (1992): "Wesley Salmon's Process Theory of Causality and the Conserved Quantity Theory", en: *Philosophy of Science*, 59, pp. 195-216.
- Dowe, P. (1995): "Causality and Conserved Quantities: A Reply to Salmon", en: *Philosophy of Science*, 62, pp. 321-333.
- Friedman, M. (1974): "Explanation and Scientific Understanding", en: *The Journal of Philosophy*, 71, pp. 5-19.
- Hempel, C. G. (1965): *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science*, New York: Free P.
- Hempel, C. G. y Oppenheim, P. (1948): "Studies in the Logic of Explanation", en: *Philosophy of Science*, 15, pp. 567-579.
- Kitcher, P. (1976): "Explanation, Conjunction and Unification", en: *The Journal of Philosophy*, 73, pp. 207-212.
- Kitcher, P. (1981): "Explanatory Unification", en: *Philosophy of Science*, 48, pp. 507-531.
- Kitcher, P. (1989): "Explanatory Unification and the Causal Structure of the World", en: Kitcher, P. y Salmon, W. (eds.), *Scientific Explanation*, Minneapolis: U. Minnesota P., pp. 410-505.
- Kitcher, P. (1993): *The Advancement of Science*, Oxford: Oxford U.P.
- Kitcher, P. y Salmon, W. (1987): "Van Fraassen on Explanation", en: *The Journal of Philosophy*, 84, pp. 315-330.
- Kitcher, P. y Salmon, W. (eds.) (1989): *Scientific Explanation*, Minneapolis: U. Minnesota P. pp. 410-505.
- Lewis, D. (1973): "Causation", en: *The Journal of Philosophy*, 70, pp. 556-567.
- Lewis, D. (1986): "Causal Explanation", en: *Philosophical Papers II*, Oxford: Oxford U.P., pp. 214-240.
- Rubén, D. (1990): *Explaining Explanation*, New York: Routledge.
- Salmon, W. (1981): "Causality: Production and Propagation", en: Asquith, P. D. y Giere, R. (eds.), *PSA 1980*, Philosophy of Science Association, East Lansing MI, pp. 49-69.
- Salmon, W. (1984): *Scientific Explanation and the Causal Structure of the World*, Princeton NJ: Princeton U.P.
- Salmon, W. (1989): "Four Decades of Scientific Explanation", en: Kitcher, P. y Salmon, W. (eds.), *Scientific Explanation*, Minneapolis: U. Minnesota P., pp. 3-219.

- Salmon, W. (1992): "Scientific Explanation", en: Salmon, M. *et al.*, *Introduction to the Philosophy of Science*, Prentice Hall, Englewood Cliffs NJ, pp. 7-41.
- Salmon, W. (1994): "Causality Without Counterfactuals", en: *Philosophy of Science*, 61, pp. 297-312.
- Salmon, W. (1997): "Causality and Explanation: A Reply to Two Critiques", en: *Philosophy of Science*, 64, pp. 461-477.
- Salmon, W. (1998): *Causality and Explanation*, New York: Oxford U. P.
- Salmon, W. (2001a): "Causal *versus* Noncausal Explanation" (mecanuscrito).
- Salmon, W. (2001b): "The Structure of Causal Explanation" (mecanuscrito).
- Van Fraassen, B. (1977): "The Pragmatics of Explanation", en: *American Philosophical Quarterly*, 14, pp. 1143-1150.
- Van Fraassen, B. (1980): *The Scientific Image*, Oxford: Clarendon P.

3. Límites de la explicación como subsunción

3.1 Subsunción y ampliación teórica

La respuesta de Díez a la pregunta acerca de qué es explicar es clara y directa: explicar es subsumir el fenómeno en un patrón nómico (cf. Díez, 2002). Para articular esta respuesta, el modelo de explicación como subsunción hace uso del esquema de Van Fraassen acerca de los contextos explicativos. Es ahí donde surgen algunas consideraciones que son reveladoras al respecto de lo que está implícito en esta respuesta. Lo primero que quisiera destacar es que las nociones de *clase de contraste X* y *relación de relevancia explicativa R* son altamente indicadoras de lo que podríamos llamar la “carga teórica” en la determinación del fenómeno. Con esto lo que quiero indicar es que el fenómeno por el cual se pide una explicación no tiene una naturaleza definida independientemente del contexto teórico explicativo desde el cual se aborda. Esto es evidente cuando, en la síntesis de las ideas de Van Fraassen, Díez dice:

junto al tema debe incluirse (usualmente de modo implícito determinado por el contexto) una clase de alternativas, la *clase de contraste*, frente a las que se contraponen el hecho por cuya razón se inquiriere (Díez, 2002).

La clase de contraste es, entonces, un conjunto de “posibles cursos de acción del fenómeno”. Pero no se trata de *todas* las posibilidades. El rango de posibilidades está limitado a cursos de acción posibles teóricamente cercanos al actual. En este estricto sentido, cuando se habla de la carga teórica en la determinación del fenómeno, no se quiere sugerir de manera alguna que los fenómenos que explica la ciencia sean inventados por los teóricos (esa es una absurda idea a la que con demasiada frecuencia se le ha prestado una atención exagerada). La idea de la carga teórica en la determinación del fenómeno está para recordarnos, como lo han hecho filósofos de la ciencia de diferentes tendencias y gustos, que toda identificación de un fenómeno de estudio se hace desde una teoría. Pero, en la medida en que la caracterización de la explicación como subsunción teórica se apoya en las nociones de *tema* y *clase de contraste*, surge una consideración que no es menos importante.

[...] puede haber –sostiene Díez– varios tipos de respuesta adecuada dependiendo de qué relación se considere en ese contexto que es relevante para poder considerar a una explicación adecuada. Hasta que el contexto no determina *el tipo* de respuesta que se considera explicativa, la demanda de explicación está indeterminada (*ibid.*).

Esto quiere decir que la determinación de lo que cuenta como explicación es contextual. De manera que no sólo no hay algo así como “El” fenómeno, con

mismo hecho. Pero esto está sujeto, como claramente lo anota Díez, a que exista una relación reductiva entre las explicaciones. Es decir, se puede hacer una selección, si las explicaciones alternativas pueden ser subsumidas en una teoría reductora que, o bien elimine las diferencias, o bien muestre que sólo una de las explicaciones es subsumible en esta “metaexplicación”.

Sin embargo, la posibilidad de evaluar el poder explicativo de dos explicaciones alternativas, parece estar sujeta a la misma condición de que exista una relación reductiva entre las explicaciones. Esto porque sólo sobre la base de tal relación contaríamos con criterios comunes de evaluación del poder explicativo. Si, por otra parte, no se da tal relación reductiva, entonces no contaríamos con criterios comunes de evaluación del poder explicativo y, como dice Díez, tendríamos explicaciones desconectadas en competencia. Así, lo que tenemos es una relación muy estrecha entre las dos respuestas. Debemos considerar la explicación como una virtud de las teorías porque sobre la base de este poder explicativo es posible seleccionar teorías que ofrecen explicaciones divergentes de un mismo hecho. Pero esta selección está sujeta a que entre las explicaciones alternativas haya una relación reductiva. Estas consideraciones, sin embargo, parecen llevarnos a concluir que las respuestas ofrecidas son apropiadas sólo en el contexto de teorías en competencia que pueden ser subsumidas en metateorías o metaexplicaciones reductoras. Pero en el caso de encontrarnos con explicaciones desconectadas en competencia como, de nuevo, parece ser el caso de lo que Kuhn concibió como el momento del cambio revolucionario, las respuestas del modelo de subsunción se enfrentan con la dificultad de no poder llevar a cabo la selección teórica sobre la base de la evaluación del poder explicativo de las teorías, pues, cuando se trata de teorías desconectadas en competencia, no se cuenta con una metaexplicación reductora desde la cual hacer tal evaluación.

La conclusión general de este análisis sería, entonces, que el modelo de explicación como subsunción teórica ampliativa, aunque recoge los modos paradigmáticos de explicación científica, parece estar circunscrito a uno de los momentos del desarrollo de la actividad científica, a saber: el momento en que la ciencia cuenta con teorías conocidas que no están sujetas a revisión. Pero aún haría falta ver en detalle la pertinencia de este modelo para dar cuenta de la explicación en el momento en que la ciencia cuestiona el poder explicativo de las teorías conocidas y se aventura en la formulación de nuevas teorías para enfrentarse a lo desconocido. Si tomamos en cuenta la tesis de Putnam de que estos dos momentos determinan la dialéctica del progreso científico, entonces al modelo de explicación como subsunción le estaría faltando considerar un elemento importante en lo que tiene que explicar.

3. Límites de la explicación como subsunción

3.1 Subsunción y ampliación teórica

La respuesta de Díez a la pregunta acerca de qué es explicar es clara y directa: explicar es subsumir el fenómeno en un patrón nómico (cf. Díez, 2002). Para articular esta respuesta, el modelo de explicación como subsunción hace uso del esquema de Van Fraassen acerca de los contextos explicativos. Es ahí donde surgen algunas consideraciones que son reveladoras al respecto de lo que está implícito en esta respuesta. Lo primero que quisiera destacar es que las nociones de *clase de contraste X* y *relación de relevancia explicativa R* son altamente indicadoras de lo que podríamos llamar la “carga teórica” en la determinación del fenómeno. Con esto lo que quiero indicar es que el fenómeno por el cual se pide una explicación no tiene una naturaleza definida independientemente del contexto teórico explicativo desde el cual se aborda. Esto es evidente cuando, en la síntesis de las ideas de Van Fraassen, Díez dice:

junto al tema debe incluirse (usualmente de modo implícito determinado por el contexto) una clase de alternativas, la *clase de contraste*, frente a las que se contrasta el hecho por cuya razón se inquiere (Díez, 2002).

La clase de contraste es, entonces, un conjunto de “posibles cursos de acción del fenómeno”. Pero no se trata de *todas* las posibilidades. El rango de posibilidades está limitado a cursos de acción posibles teóricamente cercanos al actual. En este estricto sentido, cuando se habla de la carga teórica en la determinación del fenómeno, no se quiere sugerir de manera alguna que los fenómenos que explica la ciencia sean inventados por los teóricos (esa es una absurda idea a la que con demasiada frecuencia se le ha prestado una atención exagerada). La idea de la carga teórica en la determinación del fenómeno está para recordarnos, como lo han hecho filósofos de la ciencia de diferentes tendencias y gustos, que toda identificación de un fenómeno de estudio se hace desde una teoría. Pero, en la medida en que la caracterización de la explicación como subsunción teórica se apoya en las nociones de *tema* y *clase de contraste*, surge una consideración que no es menos importante.

[...] puede haber –sostiene Díez– varios tipos de respuesta adecuada dependiendo de qué relación se considere en ese contexto que es relevante para poder considerar a una explicación adecuada. Hasta que el contexto no determina *el tipo* de respuesta que se considera explicativa, la demanda de explicación está indeterminada (*ibid.*).

Esto quiere decir que la determinación de lo que cuenta como explicación es contextual. De manera que no sólo no hay algo así como “El” fenómeno, con

una naturaleza intrínseca, independiente de la teoría, sino que tampoco hay algo así como “La” explicación. Estas son consideraciones que no deberían pasar inadvertidas cuando se dice que “explicar es subsumir el fenómeno en un patrón nómico”.

Contando con una respuesta acerca de qué es explicar, veamos ahora cómo se plantea la relación entre subsunción y ampliación teórica, examinando la respuesta a la pregunta: ¿Cómo es posible la ampliación teórica? Esta es la pregunta acerca de si una teoría que explica un cierto tipo de fenómenos podría ampliar el dominio I de aplicaciones pretendidas. Si se advierte que lo que cuenta como explicación depende del contexto teórico, entonces se puede ofrecer una respuesta como la que de hecho ofrece el modelo de explicación como subsunción en los siguientes términos: la ampliación teórica es posible porque en el sistema están “presentes” constituyentes T-teóricos comportándose con los T-no teóricos como las constricciones lo establecen (cf. Díez, 2002). Esta es una respuesta a la pregunta acerca de cómo es posible la ampliación teórica, porque indica que, ante la presencia de un fenómeno que requiere explicación, el teórico procederá a hacer una descripción del fenómeno construyendo un modelo parcial del sistema en el que hará uso de expresiones que tienen términos tanto T-teóricos como T-no teóricos; y una vez construido este modelo, llevará a cabo la subsunción, es decir, mostrará que los términos T-teóricos se comportan con esta descripción del fenómeno tal como las constricciones de la teoría lo establecen. En este punto, Díez hace una alusión al tipo de análisis que estoy intentando poner de relieve. Cuando habla de la *clase de contraste X* y la *relación de relevancia explicativa R*, dice: “En algunos contextos, como los científicos en períodos de ciencia normal, [X y R] están fijados con bastante rigidez, pero en otros pueden ser muy variables” (*ibid.*).

En “The Corroboration of Theories” (1991), Hilary Putnam identifica la situación de la ciencia normal con un esquema explicativo que podría resumirse en los siguientes términos: en las condiciones de la ciencia normal, se cuenta con un hecho a explicar y con una teoría conocida que establece algunas constricciones formales, pero hace falta establecer cuáles son las hipótesis auxiliares que permitirían completar el conjunto de constricciones que determinan el fenómeno. En un intento por acercar el análisis de Van Fraassen al de Putnam podría decirse que los *temas*, es decir, las preguntas tipo ¿por qué $_?$? junto con la *clase de contraste*, es decir el conjunto de preguntas ¿por qué no β , o β' ... etc.? constituyen lo que Putnam está dispuesto a llamar (adoptando el lenguaje de Kuhn) un *enigma*. En este sentido, la ciencia normal consiste en la resolución de enigmas. Este rasgo de la actividad científica representa, según dice Putnam, la *tendencia*

explicativa a dar cuenta de los hechos sobre la base de una teoría que se considera conocida (cf. Putnam, 1991, pp. 131 y ss.). Desde este punto de vista, la tendencia explicativa de la ciencia es, en cierto sentido, acrítica o, si se quiere, dogmática. Esto es así porque la tendencia explicativa de la ciencia normal, en cuanto descifrador de enigmas, da por sentada la pertinencia de la teoría para determinar los rasgos relevantes del fenómeno y la relevancia de las explicaciones. El análisis podría hacerse en los siguientes términos: en el contexto de la ciencia normal, cuando la actividad científica está más marcada por su tendencia explicativa, la explicación (la resolución de enigmas) se da en términos de subsunción teórica. Sin embargo, en este contexto, la ampliación teórica no se refiere a la explicación de fenómenos “sueltos” sino a “grupos” de fenómenos (Díez, 2002). La ampliación no se da como la integración de un fenómeno aislado, sino como la integración de un fenómeno que se considera como un ejemplar de un “grupo” de fenómenos que comparten una cierta familiaridad (identificada teóricamente). En ese sentido, parece que estamos autorizados para concluir que en el modelo de explicación como subsunción, la ampliación teórica tiene lugar sobre un dominio de fenómenos “familiares” a la teoría. Pero si se considera un fenómeno que inicialmente parece un ejemplar del grupo de fenómenos familiares a la teoría y, posteriormente, resulta que éste se resiste a la subsunción, entonces, parece que nos quedan pocas alternativas. La primera es rechazar el fenómeno como un auténtico ejemplar del contenido empírico de la teoría, es decir, revelar el lado más dogmático de la tendencia explicativa al desear al fenómeno como digno de una explicación y así salvar la teoría. La segunda posibilidad es revisar la teoría, es decir, sustituir en la práctica científica la *tendencia explicativa* por la *tendencia crítica* –como la llama Putnam–. En cualquier caso habrían fallado dos cosas, a saber: el intento de explicación del fenómeno y la posibilidad de ampliación teórica. Es interesante notar que éste es precisamente el caso que Kuhn registró como la presencia de *anomalías* en el desarrollo de la empresa científica. La diferencia entre el enigma y la anomalía está en que la explicación del primero sólo requiere la identificación de las hipótesis auxiliares que completan las constricciones teóricas que dan cuenta del fenómeno, mientras que la explicación de la anomalía parece requerir una completa revisión de las hipótesis teóricas y, probablemente, la formulación de una nueva teoría. Pero es evidente que en este caso la explicación de la anomalía no estaría bien descrita como subsunción teórica ampliativa. Cuando Díez dice que hay contextos en los que X y R pueden ser muy variables parece referirse a este caso en que la explicación, más que subsunción, adopta la forma de una formulación de nuevas hipótesis teóricas. De manera que, cuando la explicación se da

como subsunción, aparentemente, entonces, estamos hablando de los contextos explicativos en los períodos de ciencia normal en los que X y R están fijos.

Esto quiere decir que para el modelo de explicación como subsunción debería estar claro que hay una especie de predeterminación teórica no sólo acerca del tipo de respuesta que se considera explicativa, sino del tipo de rasgo del fenómeno que se considera *explicable* o “digno” de una explicación. Pero sobre todo, debería estar claro que la condición para que un fenómeno sea explicable depende fuertemente de que la teoría se considere como conocida, es decir, como una teoría que no requiere revisión. Lo que me interesa anotar es que las respuestas de Díez acerca de qué es explicar y de cómo es posible la ampliación teórica son respuestas satisfactorias, pero parecen ser respuestas circunscritas a los períodos de ciencia normal. Pero en los contextos de *crisis* —como los llama Kuhn—, es decir, ante la presencia de anomalías que requieren una explicación, tales respuestas son inapropiadas porque, en estos casos, los fenómenos se resisten a la subsunción.

Ahora podría intentarse introducir una distinción entre *subsunción débil* y *subsunción fuerte*, con la esperanza de mostrar que el modelo de explicación como subsunción puede dar cuenta de los casos de subsunción débil, propios de la explicación científica en los períodos de ciencia normal, así como de los casos de subsunción fuerte propios de las “revoluciones científicas”, si se entiende que, independientemente del proceso que lleve a la reformulación teórica, la explicación consistirá de todas formas en una subsunción. Por subsunción en sentido débil entenderíamos la inclusión de fenómenos que son ejemplares de un grupo de fenómenos “familiares” a la teoría. Por otra parte, por subsunción en sentido fuerte deberíamos entender la inclusión de fenómenos anómalos en una red teórica, aun si ésta es el resultado de una nueva formulación teórica. La noción de explicación como subsunción, en un sentido débil, estaría para dar cuenta de lo que Kuhn llama *resolución de enigmas*, es decir, para ampliar el dominio I de aplicaciones pretendidas cercanas al patrón nómico. Por su parte, la noción de explicación como subsunción fuerte, tendría como finalidad dar cuenta de la forma en la cual la ciencia amplía su dominio de explicación a aplicaciones no-pretendidas o, mejor, no-previstas, lo que supongo que se llamaría *ampliación teórica fuerte*. Pero aun con esta jugada, si con el concepto de ampliación teórica fuerte se buscara ir más allá de la idea de “inclusión de fenómenos familiares en teorías conocidas”, es decir, si se quisieran abarcar los casos de “inclusión de nuevos fenómenos o anomalías en nuevas teorías”, tendríamos que el modelo de subsunción encontraría un límite. Esto porque, estrictamente hablando, si hay ampliación teórica fuerte (en el sentido de explicación de una anomalía),

entonces la explicación no se da como subsunción sino como reformulación teórica. Por otra parte, si fuera posible la subsunción, entonces no se trataría realmente de una ampliación teórica fuerte en el sentido de inclusión de un fenómeno anómalo a la red de constricciones que establece la teoría conocida. Así, si este análisis es correcto, habría unos límites claros al modelo de explicación como subsunción que estarían marcados justo en el punto en el que la ciencia normal se enfrenta a la explicación de las anomalías.

3.2. Selección teórica y poder explicativo

Ahora veamos brevemente lo que pasa con la otra pregunta básica de nuestro análisis. ¿Por qué es una virtud la explicación? Lo primero que debería advertir es que la ponencia del profesor Díez sólo trata explícitamente aspectos relacionados con el tema de la sección anterior y que los aspectos que introducen las preguntas en esta sección son, más bien, el resultado de haber acogido las preguntas de Van Fraassen para la presentación del problema. Sin embargo, creo que es posible adelantar una breve consideración sobre la forma en la cual el modelo de explicación como subsunción da cuenta de la relación entre el poder explicativo y la selección teórica.

Cuando uno se pregunta por qué la explicación es una virtud que deberíamos buscar en las teorías, la respuesta puede estar orientada por la idea básica de que “consideramos que las teorías son ‘buenas’ si explican y ‘malas’ si no lo hacen”. Sin embargo, esta es una apreciación supremamente básica y, como se sabe, en el desarrollo de la ciencia suelen presentarse situaciones que requieren un tratamiento más elaborado. Tal como lo hace notar Díez, “[...] puede haber explicaciones diferentes de un mismo hecho. Cuando ello ocurre, eso se explica si se da alguna relación reductiva entre las dos explicaciones. Si no se da, tenemos dos explicaciones desconectadas en competencia” (Díez, 2002). En este contexto es posible que uno se pregunte cómo escoger entre teorías que parecen explicar un mismo fenómeno pero que ofrecen explicaciones incompatibles o, por lo menos, divergentes. Esto es: ¿Cómo es posible la selección teórica? En este punto, parece que podemos articular la relación entre poder explicativo y selección teórica de la siguiente manera: cuando se tienen explicaciones diferentes de un mismo hecho, podríamos acudir a la virtud del *poder explicativo* como un medio para sortear la disyuntiva. De esta manera tendríamos una respuesta a nuestras preguntas en los siguientes términos: la explicación es una virtud de las teorías porque permite la selección, y, a su turno, la selección es posible evaluando el poder explicativo de las teorías en competencia. Esto es, si para algo ha de servir el hecho de que las teorías tengan la propiedad de ser explicativas es precisamente para, sobre la base de esta propiedad, escoger entre dos explicaciones de un

mismo hecho. Pero esto está sujeto, como claramente lo anota Díez, a que exista una relación reductiva entre las explicaciones. Es decir, se puede hacer una selección, si las explicaciones alternativas pueden ser subsumidas en una teoría reductora que, o bien elimine las diferencias, o bien muestre que sólo una de las explicaciones es subsumible en esta "metaexplicación".

Sin embargo, la posibilidad de evaluar el poder explicativo de dos explicaciones alternativas, parece estar sujeta a la misma condición de que exista una relación reductiva entre las explicaciones. Esto porque sólo sobre la base de tal relación contaríamos con criterios comunes de evaluación del poder explicativo. Si, por otra parte, no se da tal relación reductiva, entonces no contaríamos con criterios comunes de evaluación del poder explicativo y, como dice Díez, tendríamos explicaciones desconectadas en competencia. Así, lo que tenemos es una relación muy estrecha entre las dos respuestas. Debemos considerar la explicación como una virtud de las teorías porque sobre la base de este poder explicativo es posible seleccionar teorías que ofrecen explicaciones divergentes de un mismo hecho. Pero esta selección está sujeta a que entre las explicaciones alternativas haya una relación reductiva. Estas consideraciones, sin embargo, parecen llevarnos a concluir que las respuestas ofrecidas son apropiadas sólo en el contexto de teorías en competencia que pueden ser subsumidas en metateorías o metaexplicaciones reductoras. Pero en el caso de encontrarnos con explicaciones desconectadas en competencia como, de nuevo, parece ser el caso de lo que Kuhn concibió como el momento del cambio revolucionario, las respuestas del modelo de subsunción se enfrentan con la dificultad de no poder llevar a cabo la selección teórica sobre la base de la evaluación del poder explicativo de las teorías, pues, cuando se trata de teorías desconectadas en competencia, no se cuenta con una metaexplicación reductora desde la cual hacer tal evaluación.

La conclusión general de este análisis sería, entonces, que el modelo de explicación como subsunción teórica ampliativa, aunque recoge los modos paradigmáticos de explicación científica, parece estar circunscrito a uno de los momentos del desarrollo de la actividad científica, a saber: el momento en que la ciencia cuenta con teorías conocidas que no están sujetas a revisión. Pero aún haría falta ver en detalle la pertinencia de este modelo para dar cuenta de la explicación en el momento en que la ciencia cuestiona el poder explicativo de las teorías conocidas y se aventura en la formulación de nuevas teorías para enfrentarse a lo desconocido. Si tomamos en cuenta la tesis de Putnam de que estos dos momentos determinan la dialéctica del progreso científico, entonces al modelo de explicación como subsunción le estaría faltando considerar un elemento importante en lo que tiene que explicar.

Bibliografía

- Díez, José A. (2002): "Explicación, unificación y subsunción", en: W. González (ed.), *Diversidad de la explicación científica*, Barcelona: Ariel, pp. 73-93.
- Díez, José A. (2004): "La explicación científica: causalidad, unificación y subsunción teórica", en este volumen (pp. 383-414).
- Gaspar, Ph. (1991): "Causation and Explanation: an introductory essay", en: Boyd, Gaspar y Trout (eds.), *The Philosophy of Science*, Cambridge: MIT Press, pp. 289-297.
- Hempel, C. (1965): *La explicación científica*, Barcelona: Paidós, 1988.
- Kitcher, Ph. (1991): "Explanatory Unification", en: Boyd, Gaspar y Trout (eds.), *The Philosophy of Science* Cambridge: MIT Press, pp. 329-347.
- Putnam, H. (1991): "The Corroboration of Theories", en: Boyd, Gaspar y Trout (eds.), *The Philosophy of Science* Cambridge: MIT Press, pp. 121-137.
- Van Fraassen, B. (1991): "The Pragmatics of Explanation", en: Boyd, Gaspar y Trout (eds.), *The Philosophy of Science* Cambridge: MIT Press, pp. 317-327.

