

# ¿CUÁNDO PREGUNTAR “¿POR QUÉ?”<sup>1</sup>

## OBSERVACIONES SOBRE LA DINÁMICA DE LAS PREGUNTAS Y RESPUESTAS EN UNA INVESTIGACIÓN CIENTÍFICA

ELEONORA CRESTO

### Resumen

En este trabajo argumento a favor de la idea de que una explicación científica es una respuesta a una pregunta, aunque no necesariamente a una pregunta-por-qué. Esto no quiere decir que las preguntas-por-qué no sean elementos fundamentales de toda investigación científica: su importancia radica en que *son capaces de organizar y sistematizar un conjunto dado de creencias*. Para justificar esta afirmación, comienzo por identificar tres estadios básicos en los cuales pueden surgir preguntas-por-qué, y luego procedo a caracterizar su estructura. Muestro finalmente cómo usar este análisis para elucidar la noción de *poder de unificación* de hipótesis rivales. Entiendo que esta última noción, a su vez, constituye uno de los criterios que debemos considerar a la hora de evaluar cuál de dichas hipótesis rivales es la mejor explicación disponible, comparativamente hablando.

**PALABRAS CLAVE:** Preguntas por qué – Explicación científica – Inferencia a la Mejor Explicación – Unificación – Epistemología pragmatista

### Abstract

In this paper I take scientific explanations to be answers to questions, though not necessarily to why-questions. This, however, is not meant to imply that why-questions are not fundamental in their own peculiar way. As I understand it, why-questions are crucial elements in any research process because they are able to organize and systematize an agent's belief set. In order to justify my claim I start by identifying three basic stages in which why-questions can arise, and then I proceed to characterize their structure. I also show how to use this analysis to elucidate the notion of comparative unification power of rival hypotheses. The notion of comparative unification power, in turn, can be thought of as one of the criteria that we should take into account at the time of assessing which of such rival hypotheses is the best explanation available, comparatively speaking.

<sup>1</sup> Una versión anterior de este trabajo fue presentada en el *VII Coloquio Bariloche de Filosofía* (Bariloche, septiembre de 2006), sobre algunas ideas de mi tesis de doctorado. Agradezco especialmente a Isaac Levi, Philip Kitcher y Achille Varzi por sus comentarios y sugerencias, así como a los asistentes a dicho coloquio.

KEY WORDS: Why-questions – Scientific explanation – Inference to the best explanation – Unification – Pragmatist epistemology

## 1. Introducción

Tiempo atrás, en un artículo bien conocido, Wesley Salmon se preguntaba por qué deberíamos preguntar “por qué”.<sup>2</sup> La razón –nos decía Salmon– es que las respuestas a preguntas-por-qué generan un tipo de comprensión que no se obtiene a partir de una mera descripción de hechos. Si eludiéramos sistemáticamente el planteo de preguntas de tipo “por qué”, algo esencial estaríamos dejando afuera, ya que uno de los objetivos centrales de la ciencia es, justamente, darnos *ese* tipo de comprensión. Salmon pensaba, además, que las respuestas a preguntas-por-qué generan el tipo adecuado de comprensión porque son capaces de revelar la estructura causal del mundo.

Al menos parte de la respuesta de Salmon pertenece hoy en día al sentido común filosófico. Es difícil no estar de acuerdo con Salmon en que la búsqueda del tipo correcto de comprensión es un objetivo fundamental de la ciencia, y en que las preguntas-por-qué son igualmente fundamentales. Otros aspectos de su propuesta son un poco más problemáticos. Se ha discutido largamente, por ejemplo, si el vínculo entre explicación y causalidad es tal como Salmon lo pensó. Se ha discutido un poco menos, en cambio, el vínculo entre preguntas-por-qué y explicación.<sup>3</sup> En este trabajo argumentaré, precisamente, que la relación entre preguntas-por-qué y explicación es bastante menos lineal de lo que Salmon y otros pensaron.

Más precisamente, creo que, si bien es cierto que las preguntas-por-qué son importantes, está lejos de ser obvio en qué consiste exactamente su importancia. Según entiendo, *no* son especialmente relevantes como el *motor* de una investigación –es decir, no es cierto que toda investigación deba comenzar con una pregunta-por-qué– y tampoco creo que toda búsqueda de explicaciones involucre, de manera explícita o implícita, alguna pregunta-por-qué.

¿Por qué preguntar “por qué”, pues? Tal vez encontremos una respuesta satisfactoria a esta pregunta si intentamos contestar primero otra, íntimamente relacionada: ¿cuándo preguntamos “por qué”, en el curso de un proceso de investigación? Argumentaré que, cuando prestamos la atención debida a esta última cuestión, observamos que *la importancia de las preguntas-por-qué radica en su capacidad para organizar y sistematizar un conjunto dado de creencias*, y así, indirectamente, en su capacidad para mostrar cuándo una explicación dada es *mejor* que otras, comparativamente hablando.

<sup>2</sup> Cf. Salmon (1978).

<sup>3</sup> Pero véase la siguiente nota de este trabajo.

En lo que sigue presupondré que un proceso de investigación científica tiene generalmente como meta la *búsqueda de la mejor explicación disponible* para cierto conjunto dado de problemas iniciales. Lo llamaré, pues, de ahora en más, “proceso de inferencia a la mejor explicación”, o, para abreviar, “proceso de IME”.<sup>4</sup> En la sección 2, me ocuparé de identificar la estructura general de la dinámica de preguntas y respuestas (de cualquier tipo) en un proceso de IME, y luego, en las secciones 3 a 5, me concentraré específicamente en el caso de las preguntas y respuestas-por-qué. En particular, sugeriré que podemos identificar tres momentos, muy diferentes entre sí, en los que pueden surgir preguntas-por-qué. A partir de este análisis espero clarificar la naturaleza de la relación entre las respuestas a preguntas por qué y el concepto de explicación científica.

## 2. Un marco general para comprender la estructura de preguntas y respuestas en una investigación

¿Cómo se pone en marcha la maquinaria de una inferencia a la mejor explicación, o IME? Es habitual sugerir que el agente comienza por identificar un problema, o problemas, los que generan cierta ruta de investigación. Para precisar esta idea, diré aquí que una investigación comienza cuando un agente  $A$  identifica un conjunto  $Q_{A,t}$  de preguntas que busca responder en  $t$ . Un agente puede sentirse compelido a construir  $Q_{A,t}$  como reacción ante cierta nueva información que acaba de adquirir (por ejemplo, una observación sorprendente), o como reacción ante información antigua que por alguna razón se le presenta ahora como problemática. Diré que las creencias plenas del agente (entre las que se incluye la información específica que promueve la construcción de  $Q_{A,t}$ ) están representadas por el conjunto  $K_{A,t}$ , o conjunto de enunciados que  $A$  acepta en  $t$ .

Adviértase que la información problemática de que dispone el agente puede causar perplejidad en el sentido de llevarlo a preguntarse *por qué* algo es el caso, pero también puede llevarlo a preguntarse *cómo* sucede lo que sucede, o *qué* es lo que hace que ocurra, entre otras posibilidades. En otras palabras, los elementos de  $Q_{A,t}$  no necesitan ser preguntas-por-qué. Por otra parte, entiendo que al tratar de responder *cualquiera* de estos tipos de preguntas, el agente busca obtener diferentes tipos de *explicaciones*.<sup>5</sup> Ahora bien, es

<sup>4</sup> Para algunas discusiones recientes sobre IME véanse Day y Kinkaid (1994), Hon y Rakover (2001), Psillos (2003), Lipton (2004), y Cresto (2006). Para la posible relación entre IME y el concepto de abducción véase la introducción de Flachs y Kakas (2000).

<sup>5</sup> Para un artículo favorable a la idea de que las explicaciones no tienen por qué surgir a partir de preguntas-por-qué, véase por ejemplo Kitcher e Immerwahr (inédito). En un sentido diferente, Achinstein también sostiene que las explicaciones no necesitan restringirse a

claro que las respuestas a preguntas-por-qué pueden representarse mediante *argumentos* de una manera en que no lo pueden hacer otros tipos de preguntas: decir que “ $x$  ocurre *porque y es el caso*” invita a pensar en términos de argumentos de manera mucho más obvia que si decimos, por ejemplo, “ $x$  ocurre *de este modo*”, “ $x$  ocurre *en circunstancias y*”, o, tal vez, “ $x$  *está compuesto por y*”, por citar sólo unos pocos patrones de respuestas posibles a preguntas de formatos variados. De todo esto se sigue que no concibo a las explicaciones básicamente como argumentos de cierto tipo –pese a lo cual, como se verá, creo que el concepto de *explicación como argumento* surge necesariamente en cierto momento en el curso de una IME; la tarea de este trabajo es tratar de determinar *cuándo*. Pero antes de ello, tratemos de identificar qué elementos contiene  $Q_{A,t}$  exactamente.

Obsérvese, en primer lugar, que las preguntas de un agente típicamente tienen *presuposiciones*. En la tradición del análisis semántico de preguntas y respuestas, el término “presuposición” refiere a aquellas proposiciones que un agente *debe* dar por sentado, por una cuestión de racionalidad, si la pregunta ha de tener sentido para él.<sup>6</sup> (En lo que sigue hablaré de proposiciones o enunciados de manera indistinta). Así, por ejemplo, preguntas como “¿por qué  $\phi$ ?,” o “cómo es posible  $\phi$ ?,” presuponen que  $\phi$  ocurre. Una pregunta como “¿qué causa  $\psi$ ?,” presupone que  $\psi$  ocurre (o que puede ocurrir, según el caso) y que hay algo que lo causa, y así. Diré que una pregunta  $q$  tiene sentido para  $A$  en  $t$  si todas las presuposiciones de  $q$  de las que  $A$  es consciente en  $t$  pertenecen al conjunto de creencias plenas de  $A$ , o  $K_{A,t}$ . Si ello no es así, la pregunta está mal

---

preguntas-por-qué (1985, p. 334) – por el contrario, las explicaciones son concebidas como pares ordenados de proposiciones y *actos* de explicación, donde la proposición es una respuesta adecuada a lo que llama una “pregunta de contenido” (*content-question*). Véase también Achinstein (1985).

<sup>6</sup> De acuerdo con van Fraassen (1980), por ejemplo (y siguiendo la lógica de preguntas y respuestas de Belnap y Steel (1976)) cualquier proposición que esté implicada por todas las *respuestas directas* a una pregunta dada es una presuposición de dicha pregunta – donde una respuesta directa es una proposición que proporciona suficiente información para responder exactamente la pregunta, pero no más que lo estrictamente necesario para eso (1980, sección 4.2 del capítulo 5). Cf. también Achinstein (1983), sección 4 del capítulo 2. Luego, en el contexto de un análisis más específico de las preguntas-por-qué, van Fraassen se concentra en un tipo particular de presuposición, a saber, el tópico de la pregunta (1980, sección 4.3 del capítulo 5). En una perspectiva similar, Bromberger (1966) ya había argumentado que “entenderemos por la presuposición de una pregunta-por-qué aquella que uno diría que es el caso si, al ser interrogados por la pregunta interna de la pregunta-por-qué [es decir, la pregunta que resulta de borrar el “por qué”, y que se responde por “sí” o “no”] de un modo afirmativo, uno respondería que sí, y al ser interrogados por la pregunta interna de un modo negativo, uno respondería que no” (pp. 86-87). Véase también su (1992).

formulada, en lo que respecta a  $A$ . Ahora bien, si la pregunta tiene sentido, pueden presentarse dos situaciones bien diferentes:  $A$  puede pensar que algún miembro de  $K_{A,t}$  cuenta como respuesta a  $q$ , o que ningún miembro de  $K_{A,t}$  puede responderla. En el primer caso, diremos que la pregunta está cerrada en  $t$  para  $A$ . Además, diré que una pregunta  $q$  es *legítima* para  $A$  en  $t$  si: (i)  $q$  está bien formulada para  $A$  en  $t$ ; y (ii)  $q$  no está cerrada para  $A$  en  $t$ .

Supongamos que  $q$  es legítima para  $A$  en  $t$ . En este caso  $A$  podría sentir la necesidad de buscar una respuesta para  $q$ , pero esto no es necesario ( $A$  podría no estar preparado para dedicar tiempo y esfuerzo a la empresa, por el motivo que fuere). En otras palabras, en un tiempo dado un agente  $A$  podría interesarse en responder sólo un pequeño subconjunto de sus preguntas legítimas –tal vez ninguna. Finalmente, aún suponiendo que  $A$  quiera encontrar una respuesta a  $q$  en  $t$ , no es obvio que vaya a ser capaz de sugerir respuestas posibles adecuadas. En otras palabras,  $A$  podría sólo ser capaz de proponer respuestas posibles para un subconjunto aún más pequeño de su conjunto de preguntas legítimas en  $t$ .

¿Qué hay en  $Q_{A,t}$ , pues?  $Q_{A,t}$  contiene todas las preguntas legítimas que  $A$  quiere responder en  $t$ , y que generan respuestas posibles adecuadas (i.e., hipótesis) en lo que respecta a  $A$ . Estipularé además que  $Q_{A,t}$  no puede ser vacía: si no existen preguntas que satisfagan estas restricciones, diremos simplemente que el proceso de IME no se puso en marcha.

Veamos algunas ilustraciones de estas convenciones terminológicas. Consideremos la pregunta  $q_1$ : “¿de qué están hechos los genes?”  $q_1$  no tiene sentido, por ejemplo, para los miembros de la llamada “Escuela Biométrica”, durante los primeros años del siglo veinte, puesto que rechazan una obvia presuposición de la pregunta –a saber, que los genes *existen*. En cambio, esta misma pregunta resulta perfectamente significativa para la mayoría de los científicos en los años treinta. En particular, es perfectamente significativa para Alfred Mirsky, un famoso bioquímico interesado en el estudio de las proteínas: no sólo Mirsky confía en la existencia de los genes, sino que acepta la idea de que se trata de objetos materiales de algún tipo, aunque ignora de *qué* tipo.  $q_1$  es pues legítima, en lo que respecta a Mirsky. Sin embargo, en la década del treinta Mirsky no está activamente involucrado en buscar una respuesta para  $q_1$  –está demasiado ocupado con otras líneas de investigación. La pregunta aún es significativa para él en los años setenta, pero a esta altura la considera cerrada, de modo que ha dejado de ser legítima. Para aquel entonces ya ha aceptado que los genes son fragmentos de ADN, que es hoy en día la respuesta convencional.<sup>7</sup> Consideremos, finalmente, la pregunta  $q_2$ : “¿Cómo se autorreplican los genes?” Esta pregunta resulta ciertamente significativa, por

<sup>7</sup> Pero véase la nota 13 más abajo.

ejemplo, para Theodore Avery en los años cuarenta, y disponemos de evidencia para afirmar que está ansioso por encontrar una respuesta.<sup>8</sup> Pero no tiene la menor idea de cómo hacerlo: es sencillamente incapaz de sugerir posibles explicaciones para construir un conjunto apropiado de hipótesis. Así, si bien  $q_2$  constituye una pregunta legítima que desea responder, no puede funcionar como punto de partida de un proceso de IME. En resumen, en estos ejemplos ni  $q_1$  ni  $q_2$  están en el conjunto de preguntas  $Q$  de ninguno de los agentes propuestos, respectivamente (para los tiempos sugeridos), aunque por muy diferentes razones.<sup>9</sup>

Como puede verse,  $Q_{A,t}$  se define en parte por referencia a un conjunto de respuestas o explicaciones, al que llamaré conjunto  $E_{A,t}$ . Supondré que  $E_{A,t}$  constituye un conjunto mutuamente exclusivo y exhaustivo de hipótesis explicativas rivales, compatibles con –pero no implicadas por–  $K_{A,t}$ , capaces de responder todas las preguntas de  $Q_{A,t}$  simultáneamente, y que han sobrevivido a diversos procesos de contrastación empírica. No exigiré restricciones adicionales, formales o sustantivas, que tengan que cumplir las hipótesis de  $E_{A,t}$  para ser *explicativas* para un agente. Podría objetarse que no puede desarrollarse ningún modelo interesante de IME a menos que previamente se ofrezca una teoría de la explicación. No creo que esto sea así. Entiendo que la propiedad de ser explicativa está directamente relacionada con la habilidad que una hipótesis tiene para mejorar la comprensión global del agente, o, en otros términos, con su habilidad para promover una imagen del mundo psicológicamente satisfactoria. Y, en este sentido, la tarea de decidir qué enunciados califican como explicativos, y por ende como miembros potenciales de  $E_{A,t}$ , depende de cada agente particular. Por ejemplo, un agente  $A$  podría sostener que ninguna hipótesis es verdaderamente explicativa a menos que describa un *mecanismo causal*. Pero, por cierto, no pediré que todo agente entienda a  $E_{A,t}$  de esta manera particular. Lo que cuenta como una *respuesta* a una pregunta dada, para un agente determinado, no está sujeto a escrutinio racional. Por el contrario, creo que sí es deseable, y posible, desarrollar herramientas para proveer instrucciones objetivas que lleven a identificar, en una etapa posterior, el elemento más valioso de  $E_{A,t}$ .

En otras palabras, podemos rechazar la necesidad de desarrollar restricciones objetivas acerca de qué cuenta como una explicación para un agente dado, y aún así buscar desarrollar restricciones objetivas acerca de qué cuenta

<sup>8</sup> Avery es considerado por muchos autores como el fundador (o, al menos, uno de los fundadores) de la biología molecular, por su trabajo sobre el ADN de los neumococos, junto con MacLeod y McCarty, en (1944). En la década del cuarenta, Avery y Mirsky se vieron envueltos en una feroz controversia sobre la naturaleza de la relación entre genes y ADN.

<sup>9</sup> Análizo estos ejemplos con detalle en Cresto (2008).

como *mejor explicación* (en términos relativos), desde el punto de vista del agente. Así, de acuerdo con la presente perspectiva, la mayor preocupación de una teoría de la explicación debe ser precisamente la de decir qué hipótesis explica *mejor* que otras (relativa a un conjunto  $E_{A,t}$  dado), y no la de ofrecer elucidaciones acerca de qué significa (para un agente y un tiempo dados) decir que  $H$  explica  $e$ .

La perspectiva sobre la explicación que acabo de delinear se inspira, en un sentido amplio, en una epistemología peirceana, puesto que se ocupa de discutir cómo justificar los cambios en un estado epistémico dado, en lugar de obsesionarse con qué constituye, o no constituye, un punto de partida adecuado.<sup>10</sup> Más precisamente, esta perspectiva busca decirnos cómo debería cambiar un estado epistémico dado a través del ejercicio de identificar mejores explicaciones a partir de elementos explicativos previos. En este trabajo no daré mayores detalles acerca de los posibles mecanismos de elección de elementos de  $E_{A,t}$ .<sup>11</sup> Me limitaré a señalar un aspecto de dicho mecanismo de selección que es directamente relevante para el tema que nos ocupa. A saber, entiendo que, a la hora de elegir entre un conjunto de hipótesis rivales, deben tenerse en cuenta, entre otras cosas, las diferentes virtudes epistémicas de cada una de dichas hipótesis —propiedades tales como la simplicidad, la fertilidad, la precisión, o el poder de unificación de una hipótesis. Como veremos, esta última virtud —el poder de unificación— está estrechamente relacionada con la posibilidad de que una hipótesis dada responda diversas preguntas-por-qué. De este modo, indirectamente, veremos que el concepto de mejor explicación —o mejor elemento de un conjunto pre-seleccionado de respuestas explicativas posibles— se relaciona, como quería la tradición, con el concepto de preguntas-por-qué.

### 3. ¿Cuándo preguntar “por qué”?

Existen por lo menos tres circunstancias muy diferentes en las que puede aparecer una pregunta-por-qué en el marco de un proceso de IME —y, corres-

<sup>10</sup> Según diversos autores contemporáneos, la epistemología peirceana se identifica con un modelo de investigación “creencia-duda”, cuyos orígenes pueden rastrearse hasta Peirce (1878) (“The Fixation of Belief”). El modelo creencia-duda se opondría a un modelo epistemológico cartesiano, o modelo “duda-creencia”. Según el modelo cartesiano, tradicional, debe ser posible identificar criterios para separar creencias entre aquellas que están justificadas y aquellas que no lo están. En cambio, el modelo peirceano mantendría que sólo deben justificarse los *cambios* de creencia. Véase Levi (1997), p. 4. Para una comparación entre los dos proyectos epistemológicos, véase la Introducción de Fuhrmann (1997), y Bilgrami (2004).

<sup>11</sup> Me he ocupado extensamente de este problema en otros lugares. En particular, véase Cresto (2006), capítulo 3.

pondientemente, tres circunstancias muy diferentes en las que la perspectiva tradicional de explicaciones como argumentos, o como parte de argumentos, puede tener pretensiones de validez. Pueden surgir preguntas por-qué:

- (i) Al comienzo de una investigación – es decir, entre los elementos de  $Q_{A,t}$
- (ii) En el momento de tratar de establecer *relaciones* entre hipótesis de  $E_{A,t}$  (que no son necesariamente respuestas a preguntas-por-qué) y otras proposiciones de  $K_{A,t}$ , en un sentido a determinar. Y, finalmente,
- (iii) En el momento de considerar la conveniencia de *aceptar* (es decir, de incorporar a  $K_{A,t}$ ) alguna de las hipótesis de  $E_{A,t}$ .

El contexto (i) a esta altura no reviste mayores problemas. Como ya señalé en la sección anterior, no todas las preguntas de  $Q_{A,t}$  tienen que ser preguntas-por-qué. Pero, desde luego, las preguntas-por-qué no están *excluidas* de  $Q_{A,t}$ . En principio, cualquier pregunta-por-qué que sea legítima para el agente, que el agente tenga interés en responder, y para la cual el agente sea capaz de proponer respuestas posibles, es admisible como miembro de  $Q_{A,t}$ .

El contexto (ii) es por cierto bastante más interesante. Independientemente de qué preguntas particulares contenga  $Q_{A,t}$ , podemos reflexionar si existe algún tipo de relación entre las hipótesis explicativas de  $E_{A,t}$  (que son, recordemos, respuestas posibles a las preguntas de  $Q_{A,t}$ ) y otros enunciados de  $K_{A,t}$ . Por ejemplo, supongamos que  $Q_{A,t}$  es el conjunto de preguntas que podemos atribuir a Theodore Avery a mediados de los años cuarenta;<sup>12</sup>  $Q_{A,t}$  sin dudas contiene la pregunta “¿de qué sustancia están hechos los genes?”, y  $E_{A,t}$  contiene, entre otras cosas, la hipótesis  $H_j$ : “Los genes están hechos de ADN.”<sup>13</sup>

<sup>12</sup>  $Q_{A,t}$  es, digamos, el conjunto de preguntas que motiva la investigación que culmina con el justificadamente famoso artículo de Avery, MacLeod y McCarty (1944).

<sup>13</sup> Estrictamente hablando, podría argumentarse que la afirmación de que los genes son (o están hechos de) ADN constituye un error categorial. Éste sería el caso, por ejemplo, si creyéramos que “genes” refiere a entidades abstractas, lo cual era usual en la genética clásica de la primera mitad del siglo XX. Una discusión interesante de este punto se encuentra en Hotchkiss (1966). En cierto sentido, la idea de gen como entidad abstracta, no física, sobrevive en la biología contemporánea, en cuanto no hay maneras privilegiadas de contar segmentos de ADN. Véase Kitcher (1992) y Maienschein (1992). De cualquier manera, la expresión en cuestión (que los genes están hechos de ADN) puede considerarse como una simplificación relativamente inocua, y es por cierto ampliamente usada. Para un breve panorama de cómo evolucionó el concepto de gen a lo largo de la historia, véase Carlson (1989), pp. 259 y ss. Un buen número de consideraciones filosóficas relativas a dicha evolución pueden encontrarse en Kitcher (1982), Burian (1985), y Beurton y otros (2000).

Es claro que  $H_1$  no fue pensada originalmente como una respuesta posible a una pregunta-por-qué. Sin embargo, Avery bien puede preguntarse luego si es posible usar a  $H_1$  para explicar *por qué* se ha observado la evidencia que de hecho se ha observado (por ejemplo, por qué el llamado “principio transformador” que utiliza en sus experimentos con neumococos se depolimeriza ante la presencia de enzimas destructoras de ADN.) En otras palabras, Avery bien puede plantearse si  $H_1$  puede ser usada para responder una o varias preguntas-por-qué que involucren diferentes elementos de  $K_{A,t}$ .

Notemos que, en la medida en que  $H_1$  es capaz de responder preguntas-por-qué, podemos armar razonamientos (no necesariamente deductivos) en los cuales  $H_1$  aparezca como premisa, o por lo menos como una de las premisas, y aquello que se quiere explicar-por-qué (el “tópico” de la pregunta) aparezca como conclusión. Así, por ejemplo, a partir de  $H_1$ : “los genes están hechos de ADN”, junto con “el fenómeno de transformación bacteriana observado en el contexto de los experimentos con neumococos equivale a una modificación genética”, y “las enzimas destructoras de ADN depolimerizan muestras de ADN”, puede concluirse “esta muestra, que contiene un principio de transformación de neumococos, se depolimeriza ante la presencia de DNAsa”.

En este ejemplo,  $H_1$  se propone originalmente para explicar de qué están hechos los genes. Sin embargo, en la medida en que  $H_1$  es capaz de responder numerosas preguntas-por-qué adicionales a aquellas que iniciaron la investigación –preguntas que tienen por tópico enunciados ya aceptados en  $K_{A,t}$ – podemos ver que  $H_1$  es capaz de establecer diversas relaciones inferenciales con miembros de  $K_{A,t}$ . Y, en este sentido, puede decirse que  $H_1$  es capaz de *organizar* al conjunto  $K_{A,t}$  de cierta manera –más o menos eficiente, según el caso. Ciertamente, una manera posible de organizar un conjunto de enunciados es estableciendo argumentos entre sus miembros, que, luego, podríamos fijarnos si responden a unos pocos patrones argumentales. Esto último nos sugiere una posible estrategia para evaluar el *poder de unificación* (o *de sistematización*) de una hipótesis, siguiendo algunas ideas de Kitcher (1981, 1989).<sup>14</sup> Me ocuparé de esta estrategia en la última sección.

Consideremos finalmente el contexto (iii). A diferencia de (i) y (ii), (iii) es esencialmente metalingüístico. Independientemente de qué tipo de preguntas haya en  $Q_{A,t}$ , siempre podemos inquirir: ¿por qué vale la pena aceptar, digamos,  $H_1$  y no  $H_2$ ? En el marco de una IME, una respuesta adecuada a este tipo de pregunta de segundo orden debe poner de manifiesto que  $H_1$  explica mejor

<sup>14</sup> La posición de Kitcher sobre este tema puede rastrearse hasta su (1976), donde presenta por primera vez su crítica a Friedman (1974) –el precursor de la idea de explicación como unificación. Véase también Kitcher (1985). Para una discusión crítica sobre la naturaleza de la relación entre unificación y explicación, véase Morrison (2000).

que  $H_2$ , en algún sentido que nuestra teoría favorita de IME debería encargarse de determinar. Si esto es correcto, una IME puede considerarse un tipo de respuesta a una pregunta por-qué (más precisamente, a una pregunta-por-qué que demanda razones, antes que causas)<sup>15</sup> sin que esto signifique que la “mejor explicación” que se infiere sea ella misma una respuesta a una pregunta-por-qué. Antes bien, una IME puede decirse que es un tipo de respuesta a una pregunta-por-qué debido a que la evaluación de cierto enunciado como una “mejor explicación” es la explicación de por qué lo inferimos.

Como ya he anticipado, entiendo que la evaluación de qué cuenta como una mejor explicación no puede dejar de tomar en consideración el grado en el cual las hipótesis rivales relevantes exhiben diversas virtudes epistémicas, entre las que se cuenta su poder de unificación, o la capacidad de las hipótesis dadas para aumentar la cohesión de  $K_{A,t}$ . Y para ello evaluaremos el grado en el cual una respuesta dada a una pregunta de  $Q_{A,t}$  (que no necesita ser una pregunta por-qué) puede, a su vez, responder nuevas preguntas-por-qué que permitan establecer relaciones entre miembros de  $K_{A,t}$ .

En la sección siguiente pondré un aparato formal que pueda usarse para recoger estas sugerencias. Comenzaré por definir el concepto de una “respuesta-por-qué”, o, en una terminología tal vez más apropiada, una “explicación-por-qué” –lo que abreviaré como *explicación-P*. Luego veremos cómo definir de manera rigurosa una pregunta-por-qué a partir del concepto de una explicación-P. Con todos estos elementos, en la sección 5 estaremos en condiciones de analizar con mayor precisión en qué consiste el poder de unificación de una hipótesis.

#### 4. La estructura de las preguntas y las respuestas por-qué

Definiré explicación-P (para un agente  $A$ , en un tiempo dado  $t$ ) como una secuencia ordenada  $\langle p_1 \& \dots \& p_n, \langle e, X \rangle \rangle$ , tal que:

- (i)  $p_1 \& \dots \& p_n$  (para  $\geq 1$ ) es un enunciado significativo para  $A$ , así como cada conyunto  $p_i$  de  $p_1 \& \dots \& p_n$ .
- (ii)  $e \in K_{A,t}$ .
- (iii)  $X = \{x_p, \dots, x_k\}$  es un conjunto de enunciados significativos para  $A$ , tal que todas las negaciones de miembros de  $X$  están en  $K_{A,t}$ .
- (iv) El agente considera a  $p_1, \dots, p_n$  como premisas de un argumento (no necesariamente deductivo) cuya conclusión es  $e \& \sim x_1 \& \dots \& \sim x_k$ .
- (v) El agente considera que  $p_1 \& \dots \& p_n$  explican  $e$ .

<sup>15</sup> Agradezco esta observación a Ezequiel Zerbudis.

Este análisis (y las definiciones que siguen más abajo) está obviamente inspirado en el que ofrecen van Fraassen (1980) y Kitcher (1989), aunque, como podrá observarse, la presente propuesta difiere de la de ambos autores en varios aspectos cruciales.<sup>16</sup>

Siguiendo una terminología usual, diré que  $e$  es el tópico de la pregunta (en palabras de Peirce, el “hecho sorprendente” que el agente busca explicar), mientras que  $X$  es una “clase de contraste” para  $e$ : en el momento de pedir información sobre  $e$  podemos tener diferentes clases de contraste en mente - ¿preguntamos por qué  $e$  en lugar de  $x_1$  o  $y_1$ ? ¿O tal vez por qué  $e$  en lugar de  $x_2$  o  $y_2$ ? La clase de contraste especifica *el respecto* bajo el cual buscamos explicar a  $e$ .

Resulta claro que las explicaciones-P codifican un tipo particular de argumento, construido sobre la base de cierta comprensión pre-teórica de explicación (recogida en la cláusula (v); volveré sobre este punto en un momento).

Luego, sea  $P_{A,t}$  el conjunto de todas las explicaciones-P  $\langle p_1 \& \dots \& p_n, \langle e, X \rangle \rangle$  tales que:

- (i')  $p_1 \& \dots \& p_n$  (para  $n \geq 1$ ) es un enunciado de  $K_{A,t}$ , así como cada conyunto  $p_i$  de  $p_1 \& \dots \& p_n$ .
- (vi) Si  $\langle p_1 \& \dots \& p_n, \langle e, X \rangle \rangle$  y  $\langle q_1 \& \dots \& q_n, \langle e, Y \rangle \rangle$  están ambos en  $P_{A,t}$ , entonces o bien  $X \neq Y$ , o bien  $p_1 \& \dots \& p_n$  es lógicamente equivalente a  $q_1 \& \dots \& q_n$ , dado  $K_{A,t}$ .

$P_{A,t}$  es entonces el conjunto de las explicaciones-P que el agente  $A$  acepta en  $t$ , relativo a  $K_{A,t}$ .  $P_{A,t}$  está formalmente restringida por el postulado (vi), que especifica una condición de coherencia. Puede haber maneras genuinamente diferentes de explicar  $e$  en  $P_{A,t}$ , siempre y cuando  $e$  pertenezca a diferentes clases de contraste en cada caso—esto es, siempre y cuando diferentes explicaciones se enfoquen en distintos aspectos de  $e$ . En otras palabras, (vi) exige que los agentes no acepten más que una explicación del mismo evento, en el mismo sentido, en el mismo respecto, al mismo tiempo.

También será útil contar con la siguiente definición. Sea  $P_{E,A,t}$  el conjunto de explicaciones-P correspondientes al conjunto de hipótesis  $E_{A,t}$ . Diremos que una explicación-P  $\langle p_1 \& \dots \& p_n, \langle e, X \rangle \rangle$  está en  $P_{E,A,t}$  si y sólo si:

- (i'')  $p_1 \& \dots \& p_n$  (para  $n \geq 1$ ) es un enunciado significativo para  $A$ , y hay exactamente un conyunto  $p_i$  de  $p_1 \& \dots \& p_n$  que pertenece a  $E_{A,t}$ ; el resto de los conyuntos  $p_j$  son enunciados de  $K_{A,t}$ .

<sup>16</sup> Para un análisis alternativo véase Hintikka y Halonen (1995).

Así,  $P_{E,A,t}$  es el conjunto de explicaciones-P que codifica los argumentos contruidos sobre la base de explicaciones que se hallan en el conjunto de respuestas posibles  $E_{A,t}$ .

Con estos elementos podemos definir la noción de pregunta-por-qué (o, para abreviar, “pregunta-P”), de la siguiente manera. Diremos que una pregunta-P para un agente  $A$  en  $t$  es un triple ordenado  $\langle e, X, \mathfrak{R} \rangle$ , tal que:

- (1)  $E \in K_{A,t}$
- (2)  $X = \{x_p \dots x_k\}$  es un conjunto de enunciados significativos para  $A$ , tal que las negaciones de todos los elementos de  $X$  están en  $K_{A,t}$ .
- (3)  $\mathfrak{R}$  es el conjunto de todas las explicaciones-P  $\langle p_1 \& \dots \& p_n \langle e, X \rangle \rangle$  que  $A$  es capaz de construir en  $t$ .

( $\mathfrak{R}$  podría ser vacío, en caso de que el agente sea incapaz de proponer alguna respuesta posible a la pregunta en cuestión).

Además, una *pregunta-P asociada a  $E$* , para  $A$  en  $t$ , es una pregunta-P  $\langle e, X, \mathfrak{R} \rangle$  para  $A$  en  $t$  tal que:

$$(3') \quad \mathfrak{R} = P_{E,A,t}$$

Obsérvese que las explicaciones-P del conjunto  $P_{E,A,t}$  pueden pensarse como argumentos que tratan sobre algunas de las preguntas legítimas de  $A$  en  $t$ : a saber, aquellas preguntas-P asociadas al conjunto  $E_{A,t}$ . Más precisamente, el conjunto  $P_{E,A,t}$  comprende las muchas maneras en que los elementos de  $E_{A,t}$  pueden dar respuesta a varias preguntas-P. Las explicaciones-P del conjunto  $P_{A,t}$  por otro lado, ya no responden a preguntas legítimas de  $A$  en  $t$ : como tales explicaciones ya han sido aceptadas, las preguntas que las podrían haber originado ya están cerradas.

Por la cláusula (v), es claro que el concepto general de una explicación-P, así como el de una explicación-P *aceptada* (en el conjunto  $P_{A,t}$ ), o el de una explicación-P asociada a  $E_{A,t}$ , ya presuponen cierta comprensión pre-teórica de explicación. No cualquier secuencia ordenada posible que satisfaga los requisitos (i) a (iv) (o, alternativamente, los requisitos (i'), (ii), (iii), (iv) y (vi), o los requisitos (i'')-(iv'')) será explicativa, para un agente y un tiempo dados. Nuevamente, es el agente el que debe decidir cuáles secuencias codifican explicaciones y cuáles no. No pediré que se cumplan restricciones adicionales objetivas acerca de cuáles secuencias se permiten como explicaciones-P en general, o como miembros de  $P_{A,t}$ , o de  $P_{E,A,t}$ . En particular, la relación entre  $p_1 \& \dots \& p_n$  y  $e$ , y entre  $p_1 \& \dots \& p_n$  y la negación de los miembros de  $X$  puede, pero no necesita, ser una relación deductiva. Sin embargo, en consonancia con las observaciones vertidas en la sección 2, sí pienso que habrá algo que decir

acerca de cómo debe modificarse el conjunto  $P_{A,t}$ . Veamos esto con más detalles en la sección siguiente.

## 5. Preguntas-P y el concepto de poder de unificación

Sea  $H$  un elemento de  $E_{A,t}$ . Si el agente  $A$  expandiera su conjunto de creencias plenas  $K_{A,t}$  con la hipótesis  $H$ , en el contexto de una IME, todas las secuencias  $\langle p_1 \& \dots \& H \& \dots \& p_n \langle e_i, X_j \rangle \rangle$  de  $P_{E,A,t}$  deberían agregarse a  $P_{A,t}$ . Por (vi), esto podría resultar en que explicaciones-P sean borradas del conjunto  $P_{E,A,t}$ —en caso que el agente haya sostenido previamente maneras alternativas de explicar  $e_i$  con respecto a la misma clase de contraste  $X_j$ . Incidentalmente, nótese que una contracción de  $P_{E,A,t}$ , en la manera que acabo de sugerir, no necesariamente llevará a una contracción de  $K_{A,t}$ , a menos que supongamos que  $K_{A,t}$  también contiene enunciados de segundo orden que describen relaciones de explicación-P entre varios enunciados de primer orden (en otras palabras, a menos que exijamos que los enunciados que describen los elementos de  $P_{E,A,t}$  pertenezcan ellos mismos a  $K_{A,t}$ ), lo cual no tiene por qué ocurrir.

Ahora bien, es claro que cada una de las hipótesis rivales  $H_1, \dots, H_n$  de  $E_{A,t}$ , si llegaran a ser agregadas a  $K_{A,t}$ , nos llevarían a diferentes modificaciones de  $P_{A,t}$ . Propongo entonces que comparemos los diferentes conjuntos de explicaciones-P que resultarían en cada caso, para determinar cuál de ellos constituye la mejor sistematización de los elementos de  $K_{A,t}$ . Para esto, siguiendo a Kitcher (1989), podríamos tratar de identificar un conjunto  $\mathbf{P}_{A,t}$  de patrones de argumentación instanciados en  $P_{A,t}$ —donde, nuevamente, la carga de determinar qué patrones instancian los argumentos en  $P_{A,t}$  la tiene el agente. Intuitivamente, a veces dichos patrones serán tan sólo generalizaciones de explicaciones-P particulares que ya pertenecen a  $P_{A,t}$ . A título de ilustración, supongamos que cierto agente  $M$  (Mirsky, por ejemplo) tiene la creencia antecedente de que las proteínas cumplen un papel relevante en diversos procesos biológicos cruciales, y supongamos que piensa que esta hipótesis explica por qué muchos resultados empíricos son como son. Supongamos también que  $M$  está en duda acerca de si las proteínas también intervienen de manera esencial en la codificación de la herencia. En caso de que  $M$  aceptara esta última idea, nuevos argumentos explicativos se agregarían a  $P_{M,t}$  (donde  $t$  refiere a algún momento de la década del cuarenta). Más aún, muchos de los argumentos explicativos de  $P_{M,t}$  se convertirían ahora en instancias de un patrón más general que contiene como premisa esquemática la afirmación de que las proteínas desempeñan un papel fundamental, digamos, en todos los procesos biológicos de tipo  $\phi$ .

Sobre la base de estos conceptos, y siguiendo nuevamente sugerencias de Kitcher (1989), podemos elucidar la idea de *sistematización o unificación*

como una función del número de patrones en  $\mathbf{P}_{A,t}$  (cuantos menos mejor), y del tamaño de  $P_{A,t}$  (cuantos más elementos mejor). Luego podemos trasladar esta evaluación a las hipótesis de  $E_{A,t}$ . Diremos pues que *H es máximamente virtuosa en cuanto a su poder de unificación (para un agente A, en t, y con respecto a otros miembros de  $E_{A,t}$ ), si y sólo si H nos lleva a un conjunto  $P_{A,t}$  que alcanza el mejor compromiso entre el número de explicaciones-P y el número de patrones instanciados por tales explicaciones-P*.<sup>17</sup>

Volviendo al ejemplo de Mirsky, podríamos sospechar que la hipótesis de que los genes son proteínas ( $H_{\text{PROT}}$ ) promueve una mayor cohesión en su conjunto de creencias previas (en la década del cuarenta) que la hipótesis rival propuesta por Avery, según la cual los genes están hechos de ADN ( $H_{\text{ADN}}$ ). Mirsky puede luego disponer de la información de que  $H_{\text{PROT}}$  tiene mayor poder de unificación para así evaluar los méritos globales de cada una de las hipótesis rivales relevantes.<sup>18</sup>

## 6. Conclusión

En este trabajo propuse un esquema para analizar las preguntas y respuestas-por-qué, y busqué motivar la idea tradicional de que las preguntas-por-qué constituyen elementos especialmente importantes de todo proceso de investigación científica. Mi diagnóstico es que las preguntas-por-qué desempeñan un papel crucial a la hora de *sistematizar* conjuntos de creencias. En este sentido, brindan algunas de las herramientas necesarias para poder evaluar la satisfacción general que puede traer la aceptación de una hipótesis, y, de esta manera, las preguntas-por-qué contribuyen a nuestra comprensión de qué hace que una explicación sea, comparativamente hablando, mejor que otras disponibles.

<sup>17</sup> Aquí podrían agregarse varios refinamientos que sigan más de cerca la formulación original de Kitcher (1989), pero nada de esto modificaría la idea central. Por ejemplo, podríamos requerir que  $\mathbf{P}_{A,t}$  en realidad agrupe a un conjunto de  $\mathbf{P}^i$ s, tal que cada  $\mathbf{P}^i$  sea un conjunto *posible* de patrones instanciados por  $P_{A,t}$ . Luego, podríamos tratar de establecer cuál es el mejor  $\mathbf{P}^i$  que resulta de cada una de las modificaciones posibles de  $P_{A,t}$ . Sobre esta base, deberíamos decir que una hipótesis obtiene mayor poder de unificación si lleva a un conjunto  $P_{A,t}$  que consigue el mejor compromiso entre el número de explicaciones-P en el conjunto, y el número de patrones instanciados por dichas explicaciones-P, tal como son capturados por el mejor  $\mathbf{P}^i$  de  $\mathbf{P}_{A,t}$ . A la luz de esta propuesta,  $H_1$  tiene mayor poder de sistematización que  $H_2$  si el mejor  $\mathbf{P}^i$  que resulta de la modificación de  $\mathbf{P}_{A,t}$  debida a  $H_1$  es mejor que el mejor  $\mathbf{P}^i$  que resulta de la modificación de  $\mathbf{P}_{A,t}$  debida a  $H_2$ .

<sup>18</sup> Según entiendo, a la hora de juzgar los méritos globales de una hipótesis determinada, otras virtudes deberán entrar en consideración, así como una posible asignación de probabilidades subjetivas. Un estudio más detallado del mecanismo de elección de hipótesis nos llevaría demasiado lejos, más allá de los límites de este trabajo.

## Bibliografía

- Achinstein, Peter (1983). *The Nature of Explanation*. New York: Oxford University Press.
- Achinstein, Peter (1985). "The Pragmatic Character of Explanation." Reimpreso en David-Hillel Ruben (ed.) (1993), pp. 326-344.
- Avery, Oswald T.; MacLeod, Colin M.; y McCarty, Maclyn (1944). "Studies on the Chemical Nature of the Substance Inducing Transformation of Pneumococcal Types: Induction of Transformation by a Desoxyribonucleic Acid Fraction Isolated from Pneumococcus Type III." *Journal of Experimental Medicine*, **79**: 137-158. Reimpreso en Peters, James A. (ed.) (1959), pp.173-192.
- Belnap, Nuel y T. B. Steel (1976). *The Logic of Questions and Answers*. New Haven: Yale University Press.
- Bilgrami, Akeel (2004). "Skepticism and Pragmatism." En Denis McManus (ed.), *Wittgenstein and Scepticism*. London: Routledge.
- Bromberger, Sylvain (1966). "Why-Questions." En Colodny, Robert (ed.) (1966). *Mind and Cosmos*. Pittsburgh: University of Pittsburgh Press, pp. 86-111.
- Bromberger, Sylvain (1992). *On What We Know We Don't Know: Explanation, Theory, Linguistics, and How Questions Shape Them*. Chicago: Chicago University Press.
- Burian, Richard M. (1985). "On Conceptual Change in Biology: The Case of the Gene". En Depew D.J. and Weber B.H. (eds.) *Evolution at a Crossroads: The New Biology and the New Philosophy of Science*. Cambridge: The MIT Press, pp. 21-42.
- Carlson, Elof A. (1989). *The Gene: A Critical History*. Ames, Iowa: Iowa State University Press. (Primera edición: 1966).
- Cresto, Eleonora (2006). *Inferring to the Best Explanation: A Decision-Theoretic Approach*. Ph.D. Dissertation, New York: Columbia University.
- Cresto, Eleonora (2008). "In Search of the Best Explanation About the Nature of the Gene: Avery on Pneumococcal Transformation." *Studies in History and Philosophy of Science, Part C: Biological and Biomedical Sciences*. (En prensa.)
- Day, Timothy y Kincaid, Harold (1994). "Putting Inference to the Best Explanation in its Place." *Synthese* **98**: 271-295.
- Flachs, Peter y Kakas, Anthonis (eds.) (2000). *Abduction and Induction: Essays on their Relation and Integration*. Dordrecht: Kluwer.
- Friedman, Michael (1974). "Explanation and Scientific Understanding." *The Journal of Philosophy* **71**: 5-19.

- Hintikka, Jaakko y Halonen, Ilpo (1995). "Semantics and Pragmatics for Why-Questions." *Journal of Philosophy* **92**: 636-657.
- Hon, Giora y Rakover, Sam (eds.) (2001). *Explanation: Theoretical Approaches and Applications*. Dordrecht/ Boston/ London: Kluwer.
- Hotchkiss, Rolling (1966). "Gene, Transforming Principle and DNA." En Cairns, J., Stent, G. S y Watson, J.D. (eds). *Phage and the Origins of Molecular Biology*. Cold Spring Harbor, pp. 180-200.
- Kitcher, Philip (1976). "Explanation, Conjunction and Unification." *Journal of Philosophy* **73**: 207-212.
- Kitcher, Philip (1982). "Genes." *British Journal for the Philosophy of Science* **33**: 337-359.
- Kitcher, Philip (1981). "Explanatory Unification." *Philosophy of Science* **48**: 507-531.
- Kitcher, Philip (1985). "Two Approaches to Explanation". *The Journal of Philosophy* **82**: 632-39.
- Kitcher, Philip (1989). "Explanatory Unification and the Causal Structure of the World." En Kitcher, Philip and Salmon, Wesley (eds.), *Scientific Explanation. Minnesota Studies in the Philosophy of Science*, Vol. XIII. Minneapolis: University of Minnesota Press, pp. 410-505.
- Kitcher, Philip (1992). "Gene: Current Usages." En Evelyn Keller y Elizabeth Lloyd, (eds.) (1992). *Keywords in Evolutionary Biology*. Cambridge, Mass.: Harvard University Press, 1992, pp. 128-131.
- Kitcher, Philip and Immerwahr, Daniel (200?), "History And The Sciences." (inédito).
- Levi, Isaac (1997). *The Covenant of Reason*. Cambridge: Cambridge University Press.
- Lipton, Peter (2004). *Inference to the Best Explanation*. Segunda edición. London: Routledge.
- Maienschein, Jane (1992). "Gene: Historical Perspectives." En Keller and Lloyd (eds.) (1992), pp. 122-127.
- Morrison, Margaret (2000). *Unifying Scientific Theories. Physical Concepts and Mathematical Structures*. Cambridge: Cambridge University Press.
- Peirce, Charles S. (1878). "The Fixation of Belief". Reimpreso en *The Essential Peirce*, Vol. I. Editado por Nathan Houser y Christian Kloesel. Bloomington e Indianapolis: Indiana University Press, (1992), pp. 109-123
- Peters, James A. (ed.) (1959). *Classics Papers in Genetics*. Englewood Cliffs, N.J.: Prentice-Hall.
- Psillos, Statis (2003). "Inference to the Best Explanation and Bayesianism." En Fritz Stadler (ed.) *Institute of Vienna Circle Yearbook*, Vol.10. Dordrecht: Kluwer.

- Rubén, David-Hillel (ed.) (1993). *Explanation*. Oxford: Oxford University Press.
- Salmon, Wesley (1978). "Why ask 'why?' An Inquiry Concerning Scientific Explanation". Reproducido en Salmon (1998).
- Salmon, Wesley (1998). *Causality and Explanation*. New York: Oxford University Press.
- Van Fraassen, Bas (1980). *The Scientific Image*. Oxford: Clarendon Press.