

Explicaciones naturalistas de la conciencia fenoménica

Ángel García Rodríguez

La explicación de la conciencia en términos naturalistas ha sido, y aún es, un tema central y recurrente de la filosofía de la mente. Durante los últimos treinta o cuarenta años, se ha arrojado luz sobre el fenómeno a explicar, pero también han aflorado nuevas discrepancias. Así, se está de acuerdo en buscar una explicación de la fenomenología propia de la posesión de determinados estados mentales (especialmente, sensaciones, emociones o percepciones). Incluso se está de acuerdo en que tal explicación requiere apelar a propiedades fenoménicas de la experiencia consciente. Pero se discrepa seriamente sobre dos cuestiones: en primer lugar, sobre la naturaleza de las propiedades fenoménicas; en segundo lugar, sobre la manera de explicarlas en términos naturalistas. La primera es una discrepancia ontológica acerca de cómo es el mundo; en particular, si todas las propiedades del mundo, incluidas las propiedades fenoménicas, son propiedades físicas (fisicismo) o no. La segunda es una discrepancia epistémica, acerca de la explicación naturalista de esas propiedades. Pero ambas discrepancias están relacionadas, pues el fisicismo es una forma de naturalismo, aunque quizás no la única (véase la sección VI). En este sentido, una buena vía de aproximación al debate contemporáneo sobre la conciencia consiste en clarificar la noción de explicación naturalista subyacente al mismo (o incluso, de las distintas maneras de entender la idea misma de explicación naturalista). En las páginas siguientes se ofrece un mapa conceptual a gran escala de algunas posiciones centrales del debate contemporáneo (éste es demasiado complejo para intentar hacer justicia a todas), que sirve para introducir, al tiempo, los distintos artículos incluidos en este número.

I

Es posible que llegue el día en que todos los procesos que suceden en el cerebro cuando pensamos o nos enamoramos se describan en su totalidad en virtud de términos fisiológicos como materia y energía. Podría tratarse de una explicación completa en su propio ámbito; pero con ello no se habrá descrito, ni mucho

menos explicado, la experiencia de pensar, o de enamorarse. Esto se aplica en igual medida a las sensaciones simples. Al tener la sensación de rojo, luz de una longitud de onda determinada estimula un cierto tipo de célula en la retina, y así da inicio a impulsos nerviosos dirigidos hacia determinados centros del cerebro. Pero todo el conocimiento de longitudes de onda, células en la retina y centros nerviosos no conseguirá que un ciego comprenda la singular cualidad de rojo por oposición a la de verde o azul; podemos describir y explicar la maquinaria que subyace a la sensación, pero no la sensación misma. Los procesos materiales no pueden explicar la conciencia, como tampoco la conciencia puede explicar los procesos materiales; son cualidades distintas del ser.

Podemos explicar el mecanismo físico del cerebro, al menos en términos generales; no podemos explicar de qué modo el funcionamiento del cerebro nos hace sentir o conocer, ni siquiera por qué sentimos o conocemos. Nos hemos topado con el arcano que, con una sonrisa, disuelve la completud de una explicación mecánica, puramente fisiológica, de la vida.

Este texto es un magnífico resumen del problema contemporáneo de la conciencia en relación con la concepción materialista (físicista) de la realidad. A saber, según el fisicismo, la explicación física de la realidad habría de ser una explicación completa; pero la naturaleza de la experiencia consciente se resiste a una explicación en términos físicos; luego, la explicación física de la realidad no es una explicación completa. Pero a pesar de su actualidad, no se trata de un texto de un autor contemporáneo, sino que está extraído de *The science of life*, obra escrita por H.G. Wells, Julian Huxley y G.P. Wells, y publicada por primera vez en 1931.

Más aún, ni siquiera entonces era la idea central de este texto novedosa u original, pues como dijo William James en *The principles of psychology*, “que los cerebros den lugar a una conciencia que conoce, ése es el misterio que retorna, con independencia de qué tipo de conciencia y de qué tipo de conocimiento se trate. Las sensaciones, conciencia de simples cualidades, conllevan tanto misterio como los pensamientos, conciencia de sistemas complejos.” Todavía antes, T.H. Huxley, biólogo y amigo de Darwin, escribió en *Lessons in elementary physiology*, “cómo algo tan sorprendente como un estado de conciencia se produce como resultado de la irritación del tejido nervioso es tan inexplicable como la aparición del genio cuando Aladino frota la lámpara¹.”

Sin embargo, a pesar de esta venerable historia, las discusiones contemporáneas sobre la naturaleza de la conciencia proceden directamente de un renovado interés por la relación entre conciencia y fisicismo desde aproximadamente la década de los 70 del siglo pasado.

II

Un elemento común entre los textos anteriores y algunos influyentes escritos contemporáneos es la idea de misterio. Según Thomas Nagel (1974), que un organismo es consciente o tiene experiencias significa que, *para el propio organismo*, hay algo que es “como qué” (*what it is like*) ser ese organismo; es decir, un conjunto de sensaciones, una fenomenología, propias de ser ese organismo. Nagel añade que este “carácter subjetivo de la experiencia” plantea un serio problema a las explicaciones fisicistas de la conciencia, pues éstas tienen unos ideales de objetividad (léase, superación de todo punto de vista subjetivo) que casan mal con el fenómeno a explicar. De ahí a la idea del misterio de la conciencia hay una mínima distancia, y Nagel no es el único en recorrerla.

Un prominente defensor del misterio de la conciencia es Colin McGinn (1989). Como Huxley y James entre otros, McGinn acepta la existencia de correlaciones sistemáticas entre estados del cerebro y estados de conciencia. Sin embargo, rechaza que la correlación entre cerebro y conciencia implique milagro alguno, a diferencia de lo que sucedería si apareciese un genio al frotar una lámpara. Pero de todo modos concluye que la explicación de esas correlaciones se escapa a nuestra comprensión, dado el tipo de seres que somos, con las capacidades cognitivas que de hecho tenemos. Por lo tanto, el misterio permanece.

Conviene hacer notar dos cosas. En primer lugar, ni Nagel ni McGinn (a diferencia de otros autores contemporáneos) concluyen que el fisicismo sea falso, sino más bien que no podemos explicar por qué es verdadero, ni por tanto comprenderlo totalmente. El problema sobre el que se llama la atención es, pues, epistémico. En segundo lugar, en las discusiones contemporáneas no se habla del misterio de la mente en general, sino sólo del misterio de la conciencia. Ello se debe a que se piensa que es la naturaleza de la experiencia consciente lo que plantea un problema al fisicista, mientras que otros aspectos de la mente podrían ser explicados en principio en términos físicos. Así, a veces se defiende que hay un problema especial a la hora de explicar los estados o las propiedades conscientes que no se plantea respecto de los estados o las propiedades intencionales, ya que los segundos son funcionalizables, mientras que los primeros no [Kim (2006)]. De este modo, la idea del misterio de la conciencia descansa sobre el supuesto de que hay dos grandes tipos de fenómenos o estados mentales, los intencionales (o representacionales) y los conscientes, caracterizados en términos de contenido intencional *versus* contenido fenoménico (o experiencial o cualitativo), respectivamente. Llamemos a esta tesis la tesis de la separación entre intencionalidad y fenomenología (o conciencia). [Cp. Horgan y Tienson (2002) p. 520; Graham, Horgan y Tienson (2007) p. 468.]

El artículo de GALEN STRAWSON recogido en este número, “Intencionalidad real 3 (Por qué la intencionalidad entrafña conciencia)” ofrece una crítica de la tesis de la separación. Su argumento puede resumirse escuetamente del siguiente modo. Primero, la noción de contenido fenoménico (experiencial, cualitativo) usada habitualmente en la filosofía de la mente contemporánea es indebidamente restrictiva, pues caracteriza sólo el contenido afectivo-sensorial de la experiencia, dejando fuera su contenido cognitivo. (Por contenido cognitivo se entiende, como ejemplo paradigmático pero no único, el carácter de la experiencia de comprensión de una expresión lingüística; por ejemplo, una oración en la propia lengua materna, frente a un idioma desconocido). Segundo, el contenido cognitivo de la experiencia es lo que caracteriza un estado intencional determinado como el estado intencional que es: por ejemplo, el estado que (versa sobre) p en lugar de q ; puesto que sin ese contenido cognitivo, sería imposible determinar cuál de los antecedentes causales de un (supuesto) estado intencional es su objeto intencional. Por tanto, sólo seres capaces de experiencias con contenido cognitivo tienen (o pueden tener) estados intencionales; es decir, la intencionalidad es un fenómeno real (no simplemente una estrategia interpretativa) sólo en seres con experiencia (o conciencia). En otras palabras, la verdadera intencionalidad *entrafña* conciencia.

Ahora bien, ¿cómo afectaría el rechazo de la tesis de la separación a la discusión sobre la naturaleza de la conciencia? Podría parecer que la tesis defendida por Galen Strawson serviría para mostrar que no hay misterio en torno a la conciencia, pues si lo intencional es explicable en términos físicos, y no hay intencionalidad sin conciencia, entonces la conciencia *ha de ser* explicable en términos físicos. Pero el rechazo de la tesis de la separación no sólo incluye la idea de que la intencionalidad entrafña conciencia, sino que a veces se añade que la conciencia es conciencia intencional, es decir, que la conciencia entrafña intencionalidad. En este caso, si alguien está ya convencido de que hay un problema a la hora de explicar la naturaleza de la conciencia en términos físicos, y no hay conciencia sin intencionalidad, el rechazo de la tesis de la separación lleva a la ampliación del misterio de la conciencia. Así pues, el rechazo de la tesis de la separación es compatible con el misterio de la conciencia. [Cp. Horgan y Tienson (2002), p. 530.]

Pero no todo el mundo se deja seducir por el carácter misterioso de la conciencia. Así, John Searle (2000, 2004) reconoce que de momento no tenemos una explicación de las relaciones entre cerebro y conciencia, pero rechaza que no podamos llegar a tenerla en algún momento. Por otra parte, Daniel Dennett (2006) no sólo rechaza el carácter milagroso, o misterioso siquiera, de la conciencia, sino que además defiende su carácter mágico, no en el sentido de “magia de verdad”, sino en el sentido de magia real, la de los ilusionistas y prestidigitadores; y por ello concluye que “[a]l igual que la magia que se hace sobre el escenario, la ‘magia’ de la conciencia desafía toda explicación sólo si la consideramos sin más. Sólo cuando apreciamos las

formas no misteriosas en que el cerebro crea inofensivas ‘ilusiones del usuario’, empezaremos a imaginar cómo el cerebro crea la conciencia” [Dennett (2006) p. 93].

En este punto, es común hacer uso de la analogía con la polémica vitalismo *versus* mecanicismo en relación con la explicación de la vida para rechazar el carácter misterioso de la conciencia [Searle (2000), p. 180; Dennett (1991), p. 281 y (2006), p. 203]. La idea es que del mismo modo que no hace falta apelar a un misterioso *élan vital* para explicar el surgimiento de la vida, sino que ésta se puede explicar en términos físicos, la explicación de la conciencia no requiere salirse del ámbito de lo físico.

Sin embargo, esta analogía no convence a quienes defienden que no es posible explicar físicamente la conciencia. Así, David Chalmers [(1996), pp. 108-9, y (2002), p. 252] contraataca afirmando que la analogía no sirve para iluminar las discusiones contemporáneas sobre la posibilidad de explicar físicamente la conciencia. Pues en la polémica sobre la explicación de la vida ambas partes aceptaban que lo que se había de explicar es una serie de funciones como la adaptación, el crecimiento o la reproducción, mientras que los críticos actuales del fisicismo rechazan que explicar la experiencia consciente consista en explicar una serie de funciones cognitivas.

La previsible réplica a Chalmers aduce que es sólo la ignorancia actual respecto a los detalles de la explicación física de la experiencia consciente lo que crea la ilusión de que se trata de una polémica radicalmente distinta de la polémica vitalismo *versus* mecanicismo. (Así, mediante otra analogía, Paul Churchland (1996) sugiere que sólo la ignorancia de su tiempo explica que Goethe rechazara la explicación de la luz en términos físicos.) Y la única manera de combatir esa ilusión, podría añadirse, es mostrando que se puede dar una explicación de la experiencia consciente en virtud de habilidades y funciones cognitivas, donde éstas no plantean problema alguno a la idea de una explicación completa en términos físicos.

Parecería, pues, que estamos ante un *impasse*, donde el crítico del fisicismo se queja de que su oponente rechaza de antemano que la discusión sobre la naturaleza de la conciencia no sea una polémica sobre funciones cognitivas, mientras que el oponente protesta que el crítico del fisicismo se niega a aceptar que no pueda ser sino una polémica sobre funciones cognitivas. Para ir más allá de la mera constatación de este choque de intuiciones, podría mostrarse que hay una discrepancia de fondo acerca de lo que se entiende por una explicación de la experiencia consciente. La raíz de la discrepancia no es si, en algún sentido aún por determinar, es posible dar una explicación naturalista o científica de la conciencia, pues como veremos el propio Chalmers acepta que hay una explicación tal de la conciencia. La discrepancia se plantea, más bien, acerca de qué es una explicación naturalista de la conciencia. De hecho, no se trata sólo de una discrepancia entre los fisicistas y sus críticos, pues tampoco entre los fisicistas hay unanimidad respec-

to a esta cuestión. Así, los físicos coinciden en que una explicación naturalista de la conciencia ha de ser una explicación en términos físicos, pero discrepan en torno a cómo entender esto.

III

Podría pensarse que, para que se desvanezca la sensación de misterio, una explicación de la conciencia habría de producir un *sentimiento* de satisfacción epistémica, y así la discusión podría girar en torno a si una explicación en términos físicos produce o no ese sentimiento, o si al menos produce un proto-sentimiento que sea suficiente para mostrar que el *explanans* se adecua al *explanandum* [Rowlands (2007)]. Pero parece claro que todo el mundo estaría de acuerdo en que una explicación siempre lleva consigo alguna forma de satisfacción epistémica, con lo que esta vía no permite avanzar en la caracterización de la disputa en torno a la explicación física de la conciencia. Para avanzar se ha de introducir la noción de explicación reductiva.

Joseph Levine ha caracterizado una explicación reductiva diciendo que es una explicación en la que podemos “ver que el fenómeno a reducir está determinado epistémicamente por el fenómeno reductor, es decir, podemos ver que dados los hechos citados en la reducción, las cosas han de ser como parecen en la superficie” [Levine (1993), p. 129]. Una explicación tal sería la explicación, por ejemplo, de que el agua hierve a una determinada temperatura a nivel del mar citando hechos acerca de la estructura molecular del H₂O en determinadas condiciones. La explicación es total, puesto que el fenómeno observable se sigue, sin resquicio explicativo alguno, de los hechos acerca de la estructura molecular del H₂O. No hay, por tanto, hiato explicativo alguno entre el nivel observable y el nivel molecular.

Según Levine, la ausencia de hiato explicativo está conectada con el hecho de que una explicación reductiva incluye los dos pasos siguientes: primero, se proporciona una caracterización funcional (o sea, en términos de papeles causales) del concepto de la propiedad a reducir; y segundo, se identifica el mecanismo que realiza esa tarea causal. Por ejemplo, en la explicación anterior, se caracteriza (parcialmente) el concepto de agua en virtud de su comportamiento (sustancia que borbotea) en determinadas condiciones (a una determinada temperatura al nivel del mar), y luego se identifica la estructura molecular del H₂O como el mecanismo físico que cumple ese papel causal. De esta caracterización de lo que es una explicación reductiva, que incluye un primer paso de análisis conceptual, se sigue que una explicación reductiva establece una conexión conceptual entre el nivel del fenómeno a reducir (la ebullición del agua) y el nivel del fenómeno reductor (el comportamiento de las moléculas de H₂O).

Siguiendo la misma pauta, Chalmers (1996) ha precisado la idea de una explicación reductiva mediante la noción de superveniencia lógica o conceptual del modo siguiente: “un fenómeno natural es explicable reductivamente en virtud de propiedades de un nivel inferior si la propiedad de instanciar ese fenómeno superviene lógicamente de manera global a las propiedades de nivel inferior mencionadas” [Chalmers (1996), p. 48]. A su vez, la noción de superveniencia lógica o conceptual se caracteriza así: “[l]as propiedades B supervienen *lógicamente* a las propiedades A si ningún par de situaciones *lógicamente posibles* son idénticas respecto de sus propiedades A pero distintas respecto de sus propiedades B” [p. 35]. La idea central es, pues, que una explicación reductiva relaciona propiedades de nivel superior con otras de nivel inferior, mediante conexiones conceptuales que permiten iluminar el fenómeno a explicar en virtud de otros más básicos. Con ello se muestra que un determinado fenómeno no es primitivo o arbitrario, y en la medida en que se relaciona con otros de nivel inferior, “se elimina la idea de que está ocurriendo algo ‘extra’” [p. 49]. Podría decirse que así se muestra que, en algún sentido, el fenómeno a explicar estaba ya *contenido* en el fenómeno reductor. Esto sugiere, a su vez, la idea de implicación: como añade Chalmers, “cuando las propiedades B supervienen lógicamente a las propiedades A, podemos decir que los hechos A *implican* los hechos B, donde un hecho implica otro si es lógicamente imposible que el primero sea verdadero y el segundo no” [p. 36]. Así pues, en la medida en que los hechos A implican los hechos B, si alguien conociera todos los hechos A, conocería automáticamente y sin resquicio explicativo alguno todos los hechos B, dadas sólo determinadas capacidades lógico-rationales.

Asimismo, se puede usar la anterior noción de implicación para precisar la tesis según la cual una explicación fisicista de la realidad es una explicación completa, y por tanto ofrece una explicación reductiva de cualquier otro fenómeno. Lo primero es ofrecer una caracterización del fisicismo en virtud de la idea de superveniencia lógica global, del modo siguiente: cualquier propiedad no-física superviene lógica o conceptualmente de manera global a propiedades físicas. Ahora bien, si cualquier propiedad no-física superviene lógicamente a propiedades físicas, entonces los hechos físicos implican cualquier otro hecho, y el conocimiento de todos los hechos físicos implicaría conocimiento de cualquier otro hecho, dadas sólo determinadas capacidades lógico-rationales. Por tanto, si el fisicismo es verdadero, hay una explicación reductiva de cualquier otro hecho; es decir, una explicación tal que, dadas determinadas capacidades lógico-rationales, el conocimiento de la totalidad de los hechos físicos implica de manera *a priori* el conocimiento de cualquier hecho no-físico [Jackson (1994); Chalmers (1996)]. Llamemos a esta concepción “fiscicismo *a priori*”. (Chalmers (1996) lo denomina “fiscicismo del tipo A”.)

A partir de aquí, se puede caracterizar la disputa contemporánea en torno a la conciencia como la disputa acerca de si hay tal explicación reductiva de la conciencia (léase, del carácter subjetivo de la experiencia). Es decir, si los hechos físicos implican *a priori* como qué es tener la sensación S, sentir la emoción E o percibir que *p*. Así, los críticos del fisicismo aducen contraejemplos, conceptualmente inteligibles, a la idea de que lo físico implica *a priori* lo consciente. Según el argumento concebibilista (o la versión más popular del mismo, el argumento de los zombis), es conceptualmente posible que haya seres físicos, incluso funcionalmente, idénticos a nosotros que carezcan de experiencias conscientes [Chalmers (1996)]. Según el argumento del conocimiento, toda la información física relevante disponible (por ejemplo, acerca de la percepción visual de colores) no implica conocimiento de hechos fenoménicos (por ejemplo, como qué es tener la sensación de rojo) [Jackson (1982), (1986)].

El crítico del fisicismo concluye que si no hay implicación *a priori* desde lo físico a lo consciente, entonces no es posible ofrecer una explicación reductiva de la conciencia, sino que hay, en palabras de Levine (1983), “un hiato explicativo” entre hechos físicos y hechos conscientes. Chalmers (1996) puntualiza que, como se ha señalado en el párrafo anterior, la explicación reductiva falla a la hora de dar cuenta del carácter subjetivo de la experiencia, aunque otros aspectos de la posesión de conciencia podrían explicarse reductivamente. Los segundos plantean los llamados “problemas fáciles” de la conciencia, pero es el primero, el que ya preocupaba a Huxley, James o los coautores de *The science of life*, el verdadero “problema difícil”. Esta dualidad problemas fáciles *versus* problema difícil está directamente emparentada con la distinción entre dos nociones de conciencia, una funcional y otra fenoménica [Block (1995), (2002)]. Pues, según explica Chalmers, a diferencia del problema difícil, los problemas fáciles son problemas acerca de habilidades y funciones cognitivas, pero como ya se vio arriba, en principio no hay problema alguno para explicar reductivamente hechos funcionales. Así pues, el verdadero problema para el fisicismo es la existencia de hechos fenoménico-conscientes (o simplemente, fenoménicos), aquéllos de los que hablaba Nagel.

Varios artículos de este número conectan directamente con estos aspectos de la dialéctica contemporánea. El artículo de ESA DÍAZ-LEÓN, “We are living in a material world (and I am a material girl)”, explora el concepto de fisicismo subyacente a las discusiones contemporáneas, y especialmente al argumento concebibilista (de los zombis). En este sentido, la noción clave es, como se ha apuntado, la de superveniencia lógica global. Explicar esta noción requiere explicar dos cosas: la idea de superveniencia lógica y la de superveniencia global. El artículo de Díaz-León explica cómo el fisicismo es compatible con que la relación entre niveles de propiedades incluya el contexto, con lo que requiere relaciones entre mundos posibles en su conjunto: es decir, la superveniencia *global*. Por otro lado, la noción de superveniencia *lógica*

gica se opone a la de superveniencia nomológica: es decir, si el fisicismo es verdadero, entonces las propiedades fenoménicas supervienen a las propiedades físicas (o microfísicas) en cualquier mundo lógica o conceptualmente posible, no sólo en cualquier mundo nomológicamente posible. Sin embargo, como se explica en el artículo, el fisicismo no implica el rechazo de mundos posibles que contengan propiedades no-físicas; sólo implica que cualquier mundo posible que contenga las mismas propiedades físicas que el mundo actual, y nada más (por lo tanto, un duplicado físico mínimo del mundo actual), será un duplicado *simpliciter* del mundo actual, es decir, un duplicado en cuanto a todas sus propiedades. Se obtiene, en resumen, la siguiente caracterización del fisicismo: si el fisicismo es verdadero, cualquier mundo lógica o conceptualmente posible que sea un duplicado físico mínimo del mundo actual, será un duplicado *simpliciter* del mundo actual.

Esta caracterización constituye el trasfondo del argumento de los zombis contra el fisicismo. Pues, según el defensor de este argumento antifisicista, la posibilidad lógica o conceptual de los zombis significa que puede haber duplicados físicos mínimos del mundo actual que no sean duplicados *simpliciter*, ya que aquéllos, a diferencia de lo que sucede en el mundo actual, carecerían de propiedades fenoménicas. Frente a eso, el defensor del fisicismo *a priori* ha de defender que los zombis no son lógica o conceptualmente posibles; algo que a algunos se les antoja una tarea más que difícil. Sin embargo, como se verá más adelante, no es ésta la única defensa a disposición del fisicista.

En “Physicalism and phenomenal consciousness”, ROBERT KIRK ofrece primero una caracterización del fisicismo, y a continuación se basa en ella para mostrar una vía de solución al problema planteado por la conciencia fenoménica. En cuanto a lo primero, Kirk rechaza la caracterización del fisicismo en términos epistémicos (la de Jackson y Chalmers), en favor de una caracterización en términos ontológicos. Pues, según defiende, el fisicismo no requiere que los hechos fenoménicos sean cognoscibles *a priori* a partir de los hechos físicos, sino más bien que los hechos físicos entrañen los hechos fenoménicos por razones lógico-conceptuales. Por ello, el fisicista se compromete con lo que Kirk llama “la tesis de la implicación estricta”, según la cual “un enunciado A *implica estrictamente* un enunciado B sólo en caso de que ‘A y no-B’ sea inconsistente o incoherente por razones lógicas o conceptuales”.

A continuación, Kirk explora una posible solución al problema que el argumento del conocimiento plantea al fisicismo. Como se expuso arriba, según el argumento del conocimiento, toda la información física disponible no implica *como qué* es tener un estado consciente determinado. Kirk acepta que ese argumento plantea un problema a la caracterización epistémica del fisicismo, puesto que hace plausible la idea de que el conocimiento *a priori* de hechos fenoménicos a partir del conocimiento de (todos) los hechos físicos no equivale a la comprensión total de los hechos fenoménicos, la que se da cuando se tienen las experiencias relevantes. Sin embargo, según Kirk, al

abandonarse la caracterización epistémica del fisicismo en favor de la tesis de la implicación estricta se evita ese problema. Pues el fisicista puede argumentar que si se distingue entre la idea de que tener un estado mental se caracteriza por un contenido fenoménico (“there is something it is like”), y la idea de que se caracteriza por un contenido fenoménico *concreto* (“what it is like”), entonces si se pueden establecer relaciones de implicación estricta entre hechos físicos y hechos fenoménicos, “la satisfacción de esas condiciones fijará no sólo *que* hay un contenido fenoménico para el sujeto en cuestión, sino *como qué* es ese contenido fenoménico”. (También el artículo de William Fish defiende que la separación entre estos dos aspectos puede ayudar a responder a los argumentos antifisicistas.)

En “Molecule-for-molecule duplication”, MARIAM THALOS responde al argumento concebibilista de los zombis, criticando el uso que se hace de los experimentos mentales de duplicación molecular en defensa de una concepción antifisicista de la conciencia. La crítica es doble. En primer lugar, se defiende que los experimentos mentales basados en la pura intuición, sin el tutelaje de la ciencia, pueden llevar al extravío, presentando como posibilidades metafísicas lo que son “meras” posibilidades metafísicas (léase, posibilidades sin el tutelaje de la ciencia). Sin embargo, según la autora, las posibilidades metafísicas (auténticas, querría decirse) son una especie de extensión de las posibilidades nomológicas. Como ejemplo, Thalos se pregunta si es posible la duplicación microfísica de un imán sin propiedades magnéticas, respondiendo que del mismo modo que “se aprende a juzgar la imposibilidad nomológica de un duplicado microfísico de un imán que carezca de propiedades magnéticas, también se juzgará que es metafísicamente imposible que algo que tenga las propiedades microfísicas de un imán carezca de propiedades magnéticas”. Se abre así la puerta a la consideración de la validez de las intuiciones modales; en particular, si su conocimiento tiene o no una base empírica. La opinión de Thalos es tajante: “no se puede asumir la posibilidad metafísica con independencia de la posibilidad nomológica. Si se concede esto, y por qué no habría de concederse, las cuestiones de imaginación, y en general de epistemología, modal no ocuparán un lugar central” en la polémica contemporánea sobre la naturaleza de la conciencia fenoménica. (En el artículo de David Pineda se apuntan brevemente los problemas para justificar empíricamente el conocimiento modal.)

En segundo lugar, Thalos cuestiona un presupuesto implícito en los experimentos mentales sobre duplicación molecular, según el cual la duplicación molecular requiere atender exclusivamente a microprocesos (o mejor, a nuestras intuiciones al respecto). Frente a ello, se propone una defensa de la idea de duplicación (imitación) que no es independiente de la interacción con macroprocesos, o en sus propias palabras, “de las realidades nomológico-metafísicas de nivel superior acerca de cómo los sistemas complejos se mantienen estables”.

DIMITRIS PLATCHIAS, en su artículo “Experiencing a hard problem?”, rechaza la división entre un problema difícil y otros problemas fáciles acerca de la conciencia. Con ello, rechaza también el diagnóstico de Chalmers, según el cual la imposibilidad de explicar reductivamente la experiencia consciente se debe a que no se busca explicar una función o habilidad cognitiva. Frente a esto, Platchias sugiere que el problema de la experiencia consciente concierne también a la economía cognitiva de un sujeto, pues es el problema de explicar una capacidad relacional entre un sujeto y sus estados mentales (a saber, en los términos usados por Nagel, como qué es, *para un organismo*, ser ese organismo, o tener la sensación S, sentir la emoción E o percibir que p). Por consiguiente, no hay un problema difícil, tal como se entiende habitualmente esta expresión, sobre la experiencia consciente.

La propuesta de explicación de la experiencia consciente que se defiende en el artículo se incluiría entre las concepciones HOT (*Higher-Order Thought*), según las cuales el carácter consciente de la experiencia consiste en tener representaciones de orden superior de determinados estados mentales. Así, se establece una distinción entre dos niveles distintos de representación (o pensamiento): por una parte, el nivel en el que uno es consciente de determinados estados mentales (pensamientos no introspectivos de segundo orden); por otra parte, el nivel en que uno es consciente de ser consciente de determinados estados mentales (pensamientos introspectivos de tercer orden). Platchias precisa que la explicación del carácter fenoménico de la experiencia tiene que ver con la posesión de pensamientos de segundo orden.

A veces se dice que este tipo de concepciones de la experiencia se enfrenta a una seria, si no definitiva, objeción: ¿por qué tener un pensamiento de orden superior sobre un estado mental explica que haya una experiencia con un determinado carácter fenoménico? [Cp. Carruthers (2007), p. 282]. Pero Platchias tiene la respuesta preparada. Una cosa son las propiedades fenoménicas o *qualia*, y otra ser consciente de las mismas. Asumiendo con Nagel que lo que se ha de explicar es como qué es, para un organismo, tener determinadas experiencias con un contenido fenoménico determinado, se argumenta que sólo mediante pensamientos de orden superior los estados mentales se le presentan al sujeto como experiencias con un determinado carácter fenoménico. Más aún, en la medida en que las habilidades (cognitivas) involucradas en tener pensamientos de orden superior no implican a su vez la habilidad de ser consciente, se ha ofrecido una explicación no circular del carácter fenoménico de la experiencia consciente.

IV

Tenemos, pues, una caracterización de la discusión contemporánea sobre la conciencia, según la cual la polémica versa sobre la posibilidad de ex-

plicar reductivamente la conciencia fenoménica, donde una explicación tal requiere mostrar que los hechos físicos implican *a priori* los hechos fenoménicos. Según esto, tanto los defensores como los detractores del fisicismo estarían de acuerdo en que, si el fisicismo es una teoría completa verdadera, entonces ha de haber una explicación reductiva de los hechos fenoménicos. La diferencia entre ellos estriba en que, sobre la base de los argumentos de la concebibilidad y del conocimiento, el detractor del fisicismo niega que pueda haber una explicación reductiva de los hechos fenoménicos y por lo tanto rechaza que el fisicismo sea una teoría completa verdadera. En cambio, el fisicista defiende que hay una explicación reductiva de los hechos fenoménicos, lo cual requiere a su vez refutar el argumento de los zombis y el argumento del conocimiento. Para ello, el fisicista tendrá que negar que los zombis sean realmente concebibles (lógica o conceptualmente posibles), o que exista un hiato explicativo entre los hechos físicos y los hechos fenoménicos, y defender así la existencia de relaciones conceptuales, cognoscibles *a priori*, entre lo físico y lo fenoménico.

Pero ésta es sólo parte de la discusión contemporánea. Algunos fisicistas contemporáneos aceptan que las situaciones descritas en los argumentos de los zombis y del conocimiento son lógica o conceptualmente posibles; es decir, son situaciones inteligibles sin contradicción alguna. Aceptan, por tanto, que los zombis son conceptualmente posibles, y que existe un hiato explicativo entre los hechos físicos y los hechos fenoménicos. Por consiguiente, no se ven incluidos en la caracterización anterior de la polémica contemporánea entre defensores y detractores del fisicismo. Como se expone a continuación, subyace aquí una concepción distinta del fisicismo.

En términos generales (pero también vagos), en una concepción fisicista de la realidad no hay más propiedades que las propiedades físicas. De manera menos vaga, y muy atractiva por su simplicidad, podría caracterizarse el fisicismo en términos de una relación de identidad entre propiedades mentales y (en sentido amplio) propiedades físicas. Esto puede concretarse, a su vez, en dos versiones distintas, según se defienda que las propiedades mentales son idénticas a propiedades neurofisiológicas, o que son idénticas a propiedades funcionales (propiedades caracterizadas por su papel funcional en la economía cognitiva de un organismo). Pero es común a ambas versiones que las identidades se establecen empíricamente, atendiendo a consideraciones metodológicas de simplicidad y de poder explicativo.

Según una concepción alternativa del fisicismo, lo anterior no ilumina suficientemente el *status* de las propiedades mentales. La razón radica en que cualquiera de esas dos versiones de la identidad no recoge adecuadamente la principal aportación de la concepción funcionalista de los estados mentales: a saber, que la relación entre propiedades mentales y propiedades neurofisiológicas es una relación entre propiedades de niveles distintos, donde las primeras se implementan en las segundas. Así, la identidad entre propiedades mentales y

propiedades neurofisiológicas contraviene la idea de implementación, al presuponer relaciones uno-a-uno entre lo mental y lo neurofisiológico. Por otra parte, la identidad entre propiedades mentales y propiedades funcionales simplemente ignora la idea de implementación, con el consiguiente problema de la eficacia causal de las propiedades mentales/funcionales.

Por eso, se concluye, una caracterización adecuada del fisicismo requiere de la noción de superveniencia: es decir, una relación de covariación entre conjuntos de propiedades de niveles distintos, donde las propiedades del nivel inferior (léase, neurofisiológicas) determinan necesariamente las propiedades del nivel superior (léase, mentales). En esta concepción, la necesidad relevante es conceptual; es decir, las propiedades del nivel inferior determinan las del nivel superior en cualquier mundo conceptualmente posible. Así, podría argumentarse que, como los conceptos mentales fijan su referencia a través de propiedades funcionales (es decir, propiedades caracterizadas por una tarea causal) que se implementan en propiedades físicas (neurofisiológicas), una descripción suficientemente detallada de las propiedades físicas relevantes, determinará necesariamente las propiedades mentales, en cualquier mundo posible que sea un duplicado físico mínimo del mundo actual.

Esta última concepción es el fisicismo *a priori* de la sección anterior: existen relaciones conceptuales, cognoscibles *a priori*, entre las propiedades físicas y las mentales [Jackson (1994), (2007)]. En cambio, la primera concepción se denomina “fiscicismo *a posteriori*”, puesto que las identidades pertinentes se establecen empíricamente [Block y Stalnaker (1999); McLaughlin (2007b)]. A pesar de las diferencias, ambas concepciones tienen en común que son defensas de la idea de que no hay más propiedades que las propiedades físicas; de ahí su carácter fisicista. Pero mientras que el fisicista *a posteriori* piensa que el fisicismo queda reivindicado cuando se han descubierto las identidades pertinentes, el fisicista *a priori* busca algo más: a saber, explicar cómo esas identidades se sustentan sobre nuestros conceptos mentales.

Aplicado al problema de la naturaleza de la conciencia fenoménica, hay pues dos tesis distintas: una, que existen relaciones de determinación conceptual, cognoscibles *a priori*, entre las propiedades físicas y las propiedades fenoménicas; otra, que existen relaciones de identidad, establecidas empíricamente, entre las propiedades físicas y las propiedades fenoménicas. Así, para el fisicista *a posteriori* ya no hay que refutar la posibilidad conceptual de los zombis, ni la existencia de un hiato explicativo entre los hechos físicos y los hechos fenoménicos. Bastaría con mostrar que la posibilidad conceptual de los zombis no implica que puedan existir realmente (posibilidad metafísica); o que el hiato explicativo no implica hiato ontológico. La idea de fondo es que la posibilidad conceptual de los zombis o el hiato explicativo no son un problema para el fisicista porque éstos se derivan de la separación entre conceptos físicos y conceptos fenoménicos, mientras que lo que el fisicista niega

es que haya propiedades propiedades fenoménicas distintas e independientes de las físicas.

Así, los defensores del fisicismo *a posteriori* se han esforzado por mostrar que el argumento antifisicista a partir de la posibilidad conceptual no es concluyente. El argumento antifisicista tiene la estructura siguiente: aunque en general la posibilidad conceptual de que $X \neq Y$ no entraña posibilidad metafísica, como muestra el caso de la identidad agua = H_2O , el caso de las propiedades fenoménicas es diferente. Según Kripke, concebir que agua $\neq H_2O$ no implica que realmente agua $\neq H_2O$, porque la concebible ausencia de identidad se puede explicar en virtud de los mecanismos de fijación de la referencia del término “agua”: a saber, a través de modos de presentación *contingentes* de la propiedad referida. Supongamos que tal modo de presentación incluye las características de ser incolora, llenar mares y ríos, etc. En ese caso, sería concebible que la sustancia incolora que llena mares y ríos no tuviera la estructura molecular H_2O , pero de ahí no se sigue la posibilidad real de que agua $\neq H_2O$, si efectivamente la identidad es verdadera. Sin embargo, el término “dolor” (por poner un ejemplo de término que expresa un concepto fenoménico) no fija su referencia mediante indicios contingentes, sino *necesarios*, de la propiedad referida, con lo que no se puede explicar de manera semejante que la posibilidad conceptual de que dolor \neq oscilación córtico-talámica (por poner un ejemplo de propiedad física determinada) sea una ilusión. Por lo tanto, en el caso de las propiedades fenoménicas, la posibilidad conceptual de que dolor \neq propiedad física determinada implica posibilidad metafísica [Kripke (1980)]. (Chalmers (1996) apela al bidimensionalismo para obtener la misma conclusión antifisicista.)

La respuesta del fisicista *a posteriori* ha sido que aceptar que los mecanismos de fijación de la referencia de los conceptos fenoménicos no involucren modos de presentación contingentes, no es suficiente para concluir la posibilidad metafísica de que dolor \neq propiedad física determinada. Así, se aduce que, a diferencia de los conceptos de la física, la referencia de los conceptos fenoménicos se fija de manera directa, desde la perspectiva de la primera persona, de tal modo que resulta difícil integrar los conceptos fenoménicos con los conceptos físicos, surgiendo así la posibilidad conceptual de que refieran a propiedades distintas. Sin embargo, de manera análoga al caso anterior del agua, de ahí no se sigue que realmente dolor \neq propiedad física determinada, si efectivamente la identidad es verdadera. A esta respuesta fisicista se la conoce como “la estrategia de los conceptos fenoménicos” [Loar (1997); Perry (2001), Papineau (2002)].

Hay un par de artículos de este número que consideran explícitamente el valor de esta estrategia. En “Conciencia y dualismo”, DAVID PINEDA analiza dos influyentes argumentos dualistas contemporáneos: el argumento concebibilista de Kripke-Chalmers, y el argumento del dualismo de propiedades. Pineda rechaza el argumento concebibilista de Kripke-Chalmers, porque im-

plica una semántica de los términos teóricos (como es el caso de “oscilación córtico-talámica”) que es problemática a la luz del carácter *a posteriori* de las identidades en que aparecen esos términos. La idea subyacente es que si los términos teóricos designan rígidamente mediante indicios necesarios (no contingentes) de la propiedad designada (como pensaba Kripke), entonces enunciados de reducción interteórica del tipo “ $H_2O = Q$ ”, donde Q es una propiedad mecánico-cuántica, no contarían como enunciados necesarios *a posteriori*, dado que no habría situaciones realmente posibles en que los indicios designaran a otra propiedad que la designada de hecho. Sin embargo, los enunciados de reducción interteórica parecen ser identidades necesarias *a posteriori*, con lo que algo falla en la semántica de los términos teóricos asumida por Kripke. Pineda explora distintos recursos de los que podría servirse el dualista, pero concluye que ninguno es completamente satisfactorio a la hora de explicar el carácter *a posteriori* de los enunciados de identidad interteórica dentro del marco dualista.

Esta conclusión permitiría rechazar el argumento concebibilista, pero como argumenta Pineda, el fisicista *a posteriori* (“compatibilista”, dice Pineda) todavía no ha ganado la partida. De hecho, si el fisicista acepta la estrategia de los conceptos fenoménicos se enfrenta al reto siguiente: a saber, proporcionar una explicación de los conceptos fenoménicos tal que estos tengan un grosor cognitivo mínimo que permita dar cuenta del hiato entre conceptos fenoménicos y conceptos físico-funcionales que subyace a la posibilidad conceptual de que no refieran a las mismas propiedades, sin caer en una concepción fregeana de tales conceptos. Pues según la concepción fregeana de los conceptos, estos refieren mediante un modo de presentación de una propiedad del referente, y bajo el supuesto de que conceptos fenoménicos y conceptos físico-funcionales son distintos, la concepción fregeana llevaría al dualismo de propiedades fenoménicas *versus* propiedades físico-funcionales. Según Pineda, el fisicista compatibilista ha de nadar y guardar la ropa, y esto no es fácil.

LUCA MALATESTI, en “Phenomenal ways of thinking”, repasa dos versiones de la estrategia de los conceptos fenoménicos, la de John Perry y la de David Papineau, según la cual los conceptos fenoménicos, a diferencia de los conceptos físicos, son demostrativo-introspectivos, y por tanto fijan su referencia directamente, no mediante un modo de presentación contingente. Según defiende Malatesti, las estrategias no funcionan: o bien se da una caracterización muy poco robusta de los conceptos demostrativo-introspectivos, que no permite distinguirlos de los conceptos demostrativo-perceptivos; o bien se da una caracterización excesivamente robusta, que introduce modos de presentación en los mecanismos de referencia de los conceptos demostrativo-introspectivos, con lo que se contradice la estrategia basada en la referencia directa. Con más detalle, Malatesti argumenta que decir que los conceptos demostrativo-introspectivos refieren directamente, sin

más, no permite distinguir entre dos usos distintos del término “este”, el introspectivo y el perceptivo, con lo que no se da respuesta a la intuición antifisicista de que las diferencias cuando Mary, la neurofisióloga del famoso ejemplo de Jackson, usa “este” antes y después de salir de su encierro involucran usos perceptivos *versus* usos introspectivos. Por otra parte, atribuir la singularidad de los conceptos demostrativo-introspectivos al modo en que se presentan las propiedades a las que se hace referencia mediante el concepto corre el riesgo de contradecir la idea de que la referencia es directa. Por ello, concluye Malatesti, las versiones de la estrategia de los conceptos fenoménicos que hay actualmente son inestables.

En conclusión, el fisicista *a posteriori* tiene su propia respuesta a los argumentos antifisicistas, respuesta que no requiere mostrar que hay relaciones conceptuales entre las propiedades físicas y las fenoménicas, o que los hechos físicos implican *a priori* los hechos fenoménicos. Según la noción de explicación reductiva de la sección anterior (glosada en términos de implicación *a priori*), la explicación de la conciencia fenoménica del fisicista *a posteriori* no es una explicación reductiva. Sin embargo, el fisicista *a posteriori* niega que una explicación adecuada de la conciencia fenoménica requiera relaciones conceptuales, cognoscibles *a priori*; y propone que el descubrimiento de relaciones de identidad es una manera de descifrar el misterio de la conciencia, pues con ello se muestra que no hay más propiedades o hechos que los físicos. El fisicista *a posteriori* podría incluso añadir que ésta es otra manera de entender la noción de explicación reductiva.

V

Más allá de la discusión general entre defensores y detractores del fisicismo, la explicación de la conciencia fenoménica se ha visto como un problema especial de la concepción funcionalista de los estados mentales. Así, se ha argumentado que la posibilidad lógica o conceptual de los *qualia* ausentes y los *qualia* invertidos es una seria objeción al funcionalismo. Pues, según el funcionalismo, las propiedades mentales se caracterizan por su papel funcional en la economía cognitiva de un organismo, mientras que es posible que haya seres funcionalmente idénticos, donde sólo uno de ellos tiene propiedades fenoménicas o *qualia*, o donde las propiedades fenoménicas de uno están invertidas respecto a las del otro; con lo que las propiedades fenoménicas no son propiedades funcionales.

La respuesta funcionalista a estas objeciones ha sido matizada. En general, los funcionalistas han rechazado la posibilidad lógica o conceptual de los *qualia* ausentes, pues implicaría la imposibilidad de conocer introspectivamente que se dan determinadas propiedades fenoménicas, o alternativamente llevaría a una contraintuitiva concepción epifenoménica de tales propiedades

[Shoemaker (1975)]. Sin embargo, algunos funcionalistas han aceptado la posibilidad conceptual de los *qualia* invertidos, especialmente la inversión del espectro (por ejemplo, de colores), tanto en el mismo sujeto en momentos distintos como en sujetos diferentes, negando a continuación que el funcionalista no pueda refutar esa posibilidad [Shoemaker (1975), (1982)].

La intuición que subyace a ambas posibilidades, y que es especialmente pertinaz en el caso del espectro invertido, es que las propiedades fenoménicas son cualidades intrínsecas de la experiencia. Esta intuición es parte esencial de la concepción de la percepción del realismo indirecto, según la cual hay algo, un conjunto de apariencias o sensaciones, común a los casos de percepción verídica e ilusoria. Dado que ese elemento común es independiente de la presencia o no de determinados hechos u objetos en el mundo (lo que distingue la percepción verídica de la ilusoria), no es algo caracterizable relacionadamente, sino intrínseco a la experiencia perceptiva misma. Se obtiene así una dualidad de propiedades intencionales (determinadas por la relación con el mundo) *versus* propiedades fenoménicas (determinadas intrínsecamente) en la percepción. En ese caso, dado que las relaciones causales con el entorno determinan sólo las propiedades intencionales o representacionales de la experiencia perceptiva, es inteligible que las propiedades fenoménicas de sujetos perceptivos con idénticas propiedades representacionales estén invertidas.

Tenemos, pues, que el problema para el funcionalismo es el carácter intrínseco de las propiedades fenoménicas, ya que éstas no se caracterizan por sus relaciones (por ejemplo, causales), a diferencia de las propiedades funcionales. De ahí que un contra-argumento funcionalista a la posibilidad del espectro invertido haya sido negar la dualidad propiedad intrínseca *versus* propiedad intencional. Apoyados en la transparencia de la percepción, algunos funcionalistas han defendido que las únicas propiedades de la experiencia perceptiva son las propiedades intencionales. La idea es que, al atender introspectivamente a la experiencia perceptiva, lo único que se encuentra son los objetos del mundo presentados en ella. En concreto, no se encuentra propiedad intrínseca alguna. Por tanto, el carácter fenoménico de la experiencia (en palabras de Nagel, como qué es tener esa experiencia) se agota en sus propiedades intencionales o representacionales [Harman (1990); Tye (1995)].

Podría parecer que este argumento niega la existencia de los *qualia*, con lo que ni siquiera podrían plantearse las objeciones anteriores al funcionalismo. Pero en realidad el argumento sólo niega la existencia de propiedades intrínsecas. Sin embargo, esto es suficiente para que el funcionalista evite esas objeciones, pues éstas sólo pueden plantearse si se acepta la dualidad de propiedades fenoménicas intrínsecas *versus* propiedades intencionales. Por eso, los defensores contemporáneos de la posibilidad del espectro invertido (incluidos funcionalistas como Shoemaker) ya no hablan de propiedades intrínsecas, y consideran a los *qualia* propiedades representacionales de la experiencia. Su defensa de la posibilidad lógica del espectro invertido se basa

en rechazar que el carácter fenoménico de la experiencia esté totalmente determinado por las propiedades representacionales. En cambio, lo que proponen es que “el carácter fenoménico de las experiencias de color es un aspecto de su contenido representacional, un aspecto determinado internamente que puede variar con independencia de la parte del contenido determinada externamente” [Shoemaker (2007), p. 321].

Si se acepta esta propuesta, y por tanto la posibilidad lógica o conceptual del espectro invertido (y quizás, en general, de los *qualia* invertidos), el funcionalismo no cierra el hiato explicativo, ni soluciona el problema difícil de la conciencia, pues los hechos funcionales no implican hechos fenoménicos. En concreto, al negar la posibilidad lógica de los *qualia* ausentes, el funcionalista acepta que los hechos funcionales implican que se da un carácter fenoménico; pero al aceptar la posibilidad lógica de los *qualia* invertidos, el funcionalista está negando que los hechos funcionales impliquen que se da un carácter fenoménico *concreto*.

WILLIAM FISH, en su artículo “Relationalism and the problems of consciousness”, argumenta que al comprometerse con la separación entre la idea de que hay algo que es sentir un estado mental (“there is something it is like”), y la idea de que hay algo *determinado* que es sentir un estado mental (“what it is like”), el funcionalista da alas a una concepción antifisicista de la experiencia consciente. Por eso, Fish propone complementar el funcionalismo con lo que denomina “relacionismo”: a saber, la tesis según la cual las propiedades fenoménicas son propiedades relacionales, mediante las que se da un contacto directo básico (*acquaintance*) entre un sujeto e instancias de propiedades en el mundo real. El resultado es que, si se acepta con el funcionalista que las propiedades funcionales determinan (entrañan) propiedades fenoménicas, el carácter relacional de las propiedades fenoménicas conlleva que la existencia de conexiones funcionales adecuadas son suficientes para la *presentación* de propiedades determinadas del mundo. (Nótese la relación entre esta propuesta y la de Kirk.)

Aunque Fish defiende que el relacionismo complementa al funcionalismo, podría objetarse que algunas concepciones funcionalista-representacionistas contemporáneas, a saber aquellas que aceptan que el carácter fenoménico de la experiencia está determinado totalmente por el contenido intencional (por lo tanto, versiones externistas como la de Tye), ya incorporan las supuestas ventajas del relacionismo. Sin embargo, Fish responde que según la concepción *PANIC* de Tye, el carácter fenoménico de la experiencia es contenido intencional que, *además*, está preparado para afectar la economía cognitiva del sujeto (en concreto, la producción de creencias y deseos), con lo que el contenido intencional de verde (por ejemplo) podría darse sin el carácter fenoménico de verde. Luego, la concepción *PANIC* de Tye no sufre al relacionismo. (Pero véase el artículo de Manolo Martínez, donde se argumenta que la condición *P* de la concepción *PANIC*, la aptitud para producir

creencias y deseos, es superflua, dada una concepción teleosemántica del contenido intencional.)

En “La ‘P’ de *PANIC*. Representacionalismo y fenomenología del dolor”, MANOLO MARTÍNEZ también presenta una objeción a la concepción representacionista del carácter fenoménico de la experiencia defendida por Tye, en este caso a partir de la fenomenología del dolor. Según el representacionista, el carácter fenoménico del dolor es un contenido intencional, posiblemente que hay algún tipo de malfunción o daño en alguna parte del cuerpo. Pero aceptando que se dé tal representación, el carácter fenoménico del dolor parece ser, más bien, la propiedad de ser desagradable.

Martínez argumenta que el origen del problema se halla en un elemento superfluo de la concepción *PANIC* de Tye: a saber, la aptitud para producir creencias y deseos (la “P” de *poised*). Su objetivo no es rechazar las concepciones representacionistas del carácter fenoménico de la experiencia consciente (en este caso, dolorosa), sino mostrar que una concepción representacionista corregida puede dar cuenta de la fenomenología del dolor. La clave estriba en desarrollar el representacionismo siguiendo una concepción teleosemántica del contenido, en vez de la propuesta informativo-causal defendida por Tye. Una concepción teleosemántica permite distinguir entre dos aspectos del contenido de la experiencia de dolor: la representación de daño en alguna parte del cuerpo (aspecto indicativo), y la evitación del mismo (aspecto imperativo). Al vincular la propiedad de ser desagradable a este segundo aspecto, Martínez concluye, se clarifica la fenomenología del dolor, sin salir del marco representacionista.

VI

En las secciones anteriores se ha defendido que las discusiones contemporáneas acerca de la naturaleza de la conciencia son discusiones acerca de cómo explicar la experiencia consciente dentro de un marco naturalista. Se presupone, pues, que hay un fenómeno que explicar: el carácter fenoménico de la experiencia, o las propiedades fenoménicas o *qualia*. Según esto, cuando las posiciones eliminativistas entran en el debate, a menudo el objetivo de su ataque no es tanto la existencia misma de las propiedades fenoménicas, cuanto determinadas concepciones de las mismas [Dennett (1988)]. Las discrepancias más acentuadas son, pues, epistémicas; es decir, acerca de la posibilidad de explicar la experiencia consciente en virtud de propiedades naturales. La discrepancia fundamental es entre defensores y detractores del fisicismo: es decir, entre los que defienden y los que rechazan que se pueda explicar la experiencia consciente en virtud de propiedades físicas. Pero hay una discrepancia añadida interna al fisicismo, acerca de qué es explicar físicamente la experiencia consciente.

Con el fin de completar el mapa conceptual dibujado en estas páginas, es importante señalar que también prominentes detractores del fisicismo aceptan la posibilidad de explicar la conciencia dentro de un marco naturalista. Considérese el dualismo naturalista de Chalmers (1996), (2007). Según Chalmers, argumentos como el de los zombis muestran que las propiedades fenoménicas no supervienen lógicamente a las propiedades físicas, con lo que no es posible dar una explicación reductiva (léase, en base a relaciones conceptuales cognoscibles *a priori*) de las propiedades fenoménicas. De aquí se sigue, continúa Chalmers, una ontología dualista de propiedades, pues además de las propiedades físicas, entre las propiedades fundamentales de la realidad hay propiedades fenoménicas. Existen asimismo leyes de superveniencia natural que explican cómo las propiedades fenoménicas dependen de hecho (en el mundo actual) de las propiedades físicas, y emergen a partir de ellas. Pero nada de esto implica negar que las leyes físicas tengan poder explicativo completo en su ámbito; ni tampoco que las leyes psicofísicas se salgan del marco naturalista. Las leyes psicofísicas se hallan entre las leyes naturales básicas, que explican la existencia de propiedades fundamentales requeridas para explicar fenómenos como la experiencia consciente. (Una analogía favorita de Chalmers es cómo la explicación del electromagnetismo requirió tanto propiedades fundamentales nuevas, como nuevas leyes, sin abandonar en ningún momento el marco naturalista.)

La conclusión de Chalmers es que, aunque haya hiato ontológico entre propiedades físicas y fenoménicas, e incluso hiato explicativo (en el sentido de que no hay explicación reductiva de las propiedades fenoménicas), no hay “eterno misterio”, pues al descubrir las leyes de superveniencia natural se puede esclarecer la naturaleza de las propiedades fenoménicas en un marco naturalista no-reductivo.

Tenemos, pues, distintas acepciones del término “explicación naturalista”, según se entienda como explicación reductiva (entendida, a su vez, de manera distinta por los defensores del fisicismo *a priori* y los del fisicismo *a posteriori*), o como explicación no-reductiva (como es el caso de Chalmers). El elemento común parece ser que una explicación naturalista de la experiencia consciente nos permite entender cómo las propiedades físicas y las propiedades fenoménicas están integradas en el mundo natural, con lo que ha de ser compatible con las explicaciones físicas en general. Así entendida, una explicación naturalista de la conciencia estaría adecuadamente integrada con otras ramas del saber científico, lo cual sirve para fijar un programa de investigación con unos objetivos y unos métodos determinados: a saber, la búsqueda de la base física (neurofisiológica) de la experiencia consciente.

Pero la flexibilidad del término “explicación naturalista” no cubre cualquier investigación sobre la conciencia. Por eso, antes de terminar, conviene señalar que el interés en la conciencia propio de la tradición fenomenológica iniciada por Husserl es ajeno (aunque quizás no de manera excluyente) al

afán contemporáneo por descubrir el correlato neuronal de la conciencia. Más bien, lo que interesa a la tradición fenomenológica es el estudio de la experiencia consciente en sí misma, mediante detalladas descripciones de familiares experiencias conscientes.

En “Fenomenología de la conciencia y naturalismo”, ÓSCAR GONZÁLEZ CASTÁN propone atender al análisis de Husserl del término “conciencia”, y usarlo como guía para evaluar algunas concepciones naturalistas contemporáneas sobre la experiencia consciente. Más en concreto, el objetivo principal del artículo es criticar la concepción de la conciencia de McGinn. No se trata de ofrecer una respuesta a la tesis misterianista de McGinn (es cognitivamente imposible para seres como nosotros comprender la solución al misterio de la conciencia), sino la noción misma de conciencia que subyace a esa tesis. La crítica fundamental de González Castán es que McGinn concibe la relación entre sujeto, introspección y estado consciente, según el modelo de la percepción, lo cual es incompatible con las descripciones proporcionadas por Husserl. Así, González Castán se opone a las tesis siguientes de McGinn. Primero, que lo mismo que puede haber estados del mundo no percibibles, hay un nivel de conciencia inaccesible a la introspección (léase, no consciente). Segundo, que la relación entre introspección y estado consciente es exclusivamente causal, no intencional. Tercero, que lo mismo que en una concepción mínimamente realista la existencia de estados (percibibles) del mundo no requiere la existencia de sujetos que los perciben (o incluso que sean capaces de percibirlos), puede haber estados de conciencia sin un sujeto de los mismos. Por todo ello, concluye González Castán, la concepción de la conciencia de McGinn ha de ser rechazada.

El artículo de González Castán incluye también una sugerencia, más o menos velada, de mayor calado: la investigación contemporánea sobre la conciencia haría bien en volver la vista a Husserl y la tradición que él comenzó, y tomar nota de su enfoque. Estas páginas se han limitado a constatar parte de la variedad de cuestiones planteadas por el estudio de la conciencia. Y este número de *Teorema* es una contribución al mismo.²

*Departamento de Filosofía
Universidad de Murcia
E-30071 Murcia, España
E-Mail: agarcia@um.es*

NOTAS

¹ La cita de James aparece recogida en Kim [(2006), p. 221], y la de Huxley en Tye [(1995), p. 15] y en McLaughlin [(2007a), p. 431].

² Agradezco los comentarios de Paco Calvo Garzón, Esa Díaz León y Luis Valdes Villanueva, que me han ayudado a mejorar la versión final de este texto. La redacción del mismo se encuadra en el proyecto de investigación HUM2006-11603-C02-01 de la DGICYT, cofinanciado por el Gobierno de España y los Fondos FEDER.

REFERENCIAS BIBLIOGRÁFICAS

- BLOCK, N. (1995), “On a confusion about a function of consciousness”, *Behavioral and brain sciences*, vol. 18, pp. 227-47. [Versión resumida y corregida: “Concepts of consciousness”, en Chalmers, D. J. (ed.), pp. 206-19.]
- BLOCK, N. y STALNAKER, R. (1999), “Conceptual analysis, dualism and the explanatory gap”, *Philosophical review*, vol. 108, pp. 1-46.
- CARRUTHERS, P. (2007), “Higher-order theories of consciousness”, en Velmans, M. y Schneider, S. (eds.), pp. 277-86.
- CHALMERS, D.J. (1996), *The conscious mind*, New York, Oxford University Press. [Traducción castellana: *La mente consciente*, Barcelona, Gedisa, 1999. Traducción de José A. Álvarez.]
- (2002), “Consciousness and its place in nature”, en Chalmers, D.J. (ed.), pp. 247-72.
- (ed.) (2002), *Philosophy of mind: classical and contemporary readings*, New York, Oxford University Press.
- (2007), “Naturalistic dualism”, en Velmans, M. y Schneider, S. (eds.), pp. 359-68.
- CHURCHLAND, P. (1996), “The rediscovery of light”, *The journal of philosophy*, vol. 93, pp. 211-28.
- DENNETT, D.C. (1988), “Quining qualia”, en Marcel, A. y Bisiach, E. (eds.), *Consciousness in contemporary science*, Oxford, Oxford University Press, pp. 42-77. [Traducción castellana en Ezcurdia, M. y Hansberg, O. (comps.), *La naturaleza de la experiencia I*, México DF, UNAM, 2003, pp. 213-62.]
- (1991), *Consciousness explained*, Boston, MA, Little, Brown and Company. [Traducción castellana: *La conciencia explicada*, Barcelona, Paidós, 1995.]
- (2006), *Dulces sueños*, Buenos Aires, Katz editores. Traducción de Julieta Barba y Silvia Jawerbaum. [Publicado originalmente en 2005 por MIT Press.]
- GRAHAM, G., HORGAN, T. y TIENSON, J. (2007), “Consciousness and intentionality”, en Velmans, M. y Schneider, S. (eds.), pp. 468-84.
- HARMAN, G. (1990), “The intrinsic quality of experience”, en Tomberlin, J. (ed.) *Philosophical perspectives 4*, Atascadero, CA, Ridgeview, pp. 31-52. [Traducción castellana en Ezcurdia, M. y Hansberg, O. (comps.), *La naturaleza de la experiencia I*, México DF, UNAM, 2003, pp. 263-87.]
- HORGAN, T. y TIENSON, J. (2002), “The intentionality of phenomenology and the phenomenology of intentionality”, en Chalmers, D.J. (ed.), pp. 520-33.
- JACKSON, F. (1982), “Epiphenomenal qualia”, *Philosophical quarterly*, vol. 32, pp. 127-36. [Traducción castellana en Ezcurdia, M. y Hansberg, O. (comps.), *La naturaleza de la experiencia I*, México DF, UNAM, 2003, pp. 95-110.]
- (1986), “What Mary didn’t know”, *The journal of philosophy*, vol. 83, pp. 291-5. [Traducción castellana en Ezcurdia, M. y Hansberg, O. (comps.), *La naturaleza de la experiencia I*, México DF, UNAM, 2003, pp. 111-8.]

- (1994), “Finding the mind in the natural world”, en Casati, R., Smith, B. y White, G. (eds.), *Philosophy and the cognitive sciences*, Viena, Holder-Pichler-Tempsky, pp. 101-12. [Reimpreso en D.J. Chalmers (ed.), pp. 162-9.]
- (2007), “*A priori* physicalism”, en McLaughlin, B. y Cohen, J. (eds.), pp. 185-99.
- KIM, J. (2006), *Philosophy of mind* (2ª edición), Cambridge, MA, Westview Press.
- KRIPKE, S. (1980), *Naming and necessity*, Cambridge, MA, Harvard University Press. [Traducción castellana: *El nombrar y la necesidad*, México, UNAM, 1985.]
- LEVINE, J. (1983), “Materialism and *qualia*”, *Pacific philosophical quarterly*, vol. 64, pp. 354-61.
- (1993), “On leaving out what it’s like”, en Davies, M. y Humphreys, G. (eds.), *Consciousness*, Oxford, Oxford University Press, pp. 121-36.
- LOAR, B. (1997), “Phenomenal states”, en Block, N., Flanagan, O. y Güzeldere, G. (eds.), *The nature of consciousness*, Cambridge, MA, MIT Press, pp. 597-616.
- MCGINN, C. (1989), “Can we solve the mind-body problem?”, *Mind*, vol. 98, pp. 349-66. [Traducción castellana en Ezcurdia, M. y Hansberg, O. (comps.), *La naturaleza de la experiencia I*, México DF, UNAM, 2003, pp. 65-93.]
- MCLAUGHLIN, B. (2007a), “Type materialism for phenomenal consciousness”, en Velmans, M. y Schneider, S. (eds.), pp. 431-44.
- (2007b), “On the limits of *a priori* physicalism”, en McLaughlin, B. y Cohen, J. (eds.), pp. 200-23.
- MCLAUGHLIN, B. y COHEN, J. (eds.) (2007), *Contemporary debates in philosophy of mind*, Oxford, Blackwell.
- NAGEL, T. (1974), “What is it like to be a bat?”, *The philosophical review*, vol. 83, pp. 435-50. [Traducción castellana en Ezcurdia, M. y Hansberg, O. (comps.), *La naturaleza de la experiencia I*, México DF, UNAM, 2003, pp. 45-63.]
- PAPINEAU, D. (2002), *Thinking about consciousness*, Oxford, Clarendon Press.
- PERRY, J. (2001), *Knowledge, possibility and consciousness*, Cambridge, MA, MIT Press.
- ROWLANDS, M. (2007), “Mysterianism”, en Velmans, M. y Schneider, S. (eds.), pp. 335-45.
- SEARLE, J. (2000), *El misterio de la conciencia*, Barcelona, Paidós. [Publicado originalmente en 1997 por New York Review of Books.]
- (2004), *Mind: a brief introduction*, New York, Oxford University Press.
- SHOEMAKER, S. (1975), “Functionalism and *qualia*”, *Philosophical studies*, vol. 27, pp. 291-315. [Traducción castellana en Ezcurdia, M. y Hansberg, O. (comps.), *La naturaleza de la experiencia I*, México DF, UNAM, 2003, pp. 183-212.]
- (1982), “The inverted spectrum”, *The journal of philosophy*, vol. 79, pp. 357-81.
- (2007), “A case for *qualia*”, en McLaughlin, B. y Cohen, J. (eds.), pp. 319-32.
- TYE, M. (1995), *Ten problems of consciousness*, Cambridge, MA, MIT Press.
- VELMANS, M. y SCHNEIDER, S. (eds.) (2007), *The Blackwell companion to consciousness*, Oxford, Blackwell.
- WELLS, H.G., HUXLEY, J. y WELLS, G.P. (1931), *The science of life*, Londres, Cassell and Company Limited.