

LA CONCIENCIA

Robert Van Gulick, *Stanford Encyclopedia of Philosophy*, 2004

Puede que no haya ningún aspecto de la mente que sea más ordinario ni más desconcertante que la conciencia y la experiencia consciente del yo y del mundo. Se puede decir que el problema de la conciencia es el tema central de las teorías actuales de la mente. A pesar de no haya ninguna teoría compartida sobre la conciencia, hay un acuerdo muy amplio, si no universal, en que una explicación adecuada de la mente exige una comprensión clara de ella y del lugar de la conciencia en la naturaleza. Tenemos que comprender lo que es la conciencia y cómo se relaciona con otros aspectos, no conscientes, de la realidad.

1. Historia del tema

Probablemente, las preguntas sobre la conciencia son tan antiguas como los seres humanos. Las prácticas funerarias del neolítico parecen expresar creencias espirituales y suministran indicios antiguos de un pensamiento al menos rudimentariamente reflexivo acerca de la naturaleza de la conciencia humana. Igualmente se ha hallado que las culturas sin escritura adoptan alguna forma de visión espiritual o por lo menos animista que indica un alto grado de reflexión sobre la naturaleza de la percepción consciente.

A pesar de esto, ha habido quien ha argumentado que la conciencia tal como hoy la conocemos es un desarrollo histórico relativamente reciente que tuvo lugar algo después de la era homérica. Según esta opinión, los seres humanos más antiguos, incluidos los que lucharon en la guerra de Troya no tenían una experiencia de sí mismos como sujetos unificados de sus pensamientos y acciones, o, por lo menos, no de la manera como hoy día tenemos esa experiencia. Otros han afirmado que ni siquiera durante el período clásico había en griego una palabra que correspondiera a “conciencia”. Aunque los antiguos tenían muchas cosas que decir acerca de la mente, está menos claro que tuvieran conceptos especiales para lo que ahora concebimos como la conciencia ni que les interesara.

Aunque las palabras “consciente” y “conciencia” se usan hoy día de manera muy diferente, es probable que la insistencia de la Reforma en la primera como origen interno de la verdad desempeñara algún papel en el giro hacia lo interno tan característico de la moderna concepción reflexiva del yo. El Hamlet que se movía en el escenario en 1600 ya veía su mundo y a sí mismo con ojos profundamente modernos.

En el siglo XVII la conciencia se había convertido en el centro del pensamiento acerca de la mente. En efecto, desde mediados del siglo XVII hasta finales del XIX la conciencia se consideraba esencial o definitoria de lo mental. René Descartes definía la noción de pensamiento (*pensée*) en términos de la conciencia reflexiva o darse cuenta de sí mismo. En *Los Principios de la Filosofía* (1640) escribía,

Por la palabra “pensamiento” (“*pensée*”) entiendo todo aquello de lo que somos conscientes que opera en nosotros.

A finales del siglo XVII John Locke ofrecía una tesis parecida aunque ligeramente más refinada en su *Ensayo sobre el Entendimiento Humano* (1688).

No digo que no haya alma en un hombre porque no la sienta durante el sueño. Digo que no puede pensar en ningún momento, despierto o dormido, sin tener sensación de ello. Y tener sensación de ello no es necesario para nada sino para nuestros pensamientos, y para ellos es y siempre será necesario.

Locke renunció explícitamente a proponer ninguna hipótesis acerca de la base sustancial de la conciencia y su relación con la materia, pero claramente la consideraba esencial para el pensamiento y para la identidad personal.

G.W. Leibniz, contemporáneo de Locke, inspirándose posiblemente en su obra matemática sobre la diferenciación y la integración, ofreció una teoría de la mente en el *Discurso de Metafísica* (1686) que permitía que hubiera infinitos grados de conciencia y quizá incluso que hubiera algunos pensamientos inconscientes, las llamadas “*petites perceptions*”. Leibniz fue el primero en distinguir explícitamente entre percepción y apercepción; es decir, más o menos entre conciencia y autoconciencia. En la *Monadología* (1720) expuso su famosa analogía del molino para expresar su creencia de que la conciencia no puede originarse en la mera materia. En ella pedía al lector que se imaginara a alguien que caminara por dentro de

un cerebro a escala ampliada igual que si caminara por un molino y observara todas las operaciones mecánicas, que para Leibniz agotaban su naturaleza física. En ningún lugar, dice, vería semejante observador ningún pensamiento consciente.

A pesar del reconocimiento de Leibniz de la posibilidad de pensamientos inconscientes, durante la mayor parte de los dos siglos siguientes los dominios del pensamiento y la conciencia se consideraron más o menos idénticos. La psicología asociacionista, fuera la cultivada por Locke o luego por Hume en el siglo XVIII (1739) o en el XIX por James Mill (1829), se proponía descubrir los principios por los cuales los pensamientos o ideas conscientes interactúan o se afectan los unos a los otros. El hijo de James Mill, John Stuart Mill continuó la obra de su padre en la psicología asociacionista, pero concedió que las combinaciones de ideas pueden producir resultados que van más allá de sus partes mentales constituyentes, ofreciendo así un modelo de la emergencia mental (1865).

El modelo puramente asociacionista fue criticado a finales del siglo XVIII por Kant (1787), que argumentaba que una explicación adecuada de la experiencia y de la conciencia fenoménica exigía que la organización mental e intencional tuviera una estructura mucho más rica. La conciencia fenoménica, según Kant, no podría ser una simple sucesión de ideas asociadas, sino que, como poco, tiene que haber una experiencia de un yo consciente situado en un mundo objetivo estructurado con respecto al espacio, al tiempo y a la causalidad.

En el mundo anglo-americano los enfoques asociacionistas continuaron influyendo en la filosofía y la psicología hasta bien avanzado el siglo XX, mientras en la esfera alemana y europea se centró más el interés en la más amplia estructura de la experiencia que condujo en parte al estudio de la fenomenología, gracias a la obra de Edmund Husserl (1913, 1929), Martin Heidegger (1927), Maurice Merleau-Ponty (1945) y otros que extendieron el estudio de la conciencia a los ámbitos de lo social, lo corporal y lo interpersonal.

Cuando se estableció la psicología científica moderna a mediados del siglo XIX, todavía se igualaban, en lo principal, la mente y la conciencia, y los métodos introspectivos dominaron el área, como ocurre en la obra de Wilhelm Wundt (1897), Hermann von Helmholtz (1897), William James (1890) y Alfred Titchener (1901). Sin embargo, la relación de la conciencia con el cerebro siguió siendo un misterio, como lo expresan las famosas palabras de T.H. Huxley.

Que una cosa tan llamativa como un estado de conciencia resulte de la irritación de un tejido nervioso es igual de explicable que la aparición del Genio cuando Aladino frotaba la lámpara (1866).

A comienzos del siglo XX la conciencia fue eclipsada en la psicología científica, sobre todo en Estados Unidos con la aparición del conductismo (Watson, 1924; Skinner, 1953) aunque en Europa, movimientos como la psicología de la Gestalt la mantuvieron en el foco de investigación científica (Kohler, 1929; Köffka, 1935). En los años sesenta, el imperio del conductismo se debilitó con el surgir de la psicología cognitiva y su insistencia en el procesamiento de la información y en la construcción de modelos de los procesos mentales internos (Neisser, 1965; Gardiner, 1985). Pero a pesar de la insistencia renovada en la explicación de capacidades cognitivas como la memoria, la percepción y la comprensión del lenguaje, la conciencia siguió en gran medida postergada durante varias décadas más.

En los años ochenta y noventa hubo un gran resurgimiento de las investigaciones científicas y filosóficas sobre la naturaleza y las bases de la conciencia (Baars, 1988; Dennett, 1991; Penrose, 1989, 1994; Crick, 1994; Lycan 1987, 1966; Chalmers, 1996). Una vez que la conciencia hubo vuelto a ser objeto de discusión, hubo una rápida proliferación de investigaciones y una inundación de libros y artículos, así como de revistas especializadas (*The Journal of Consciousness Studies*, *Consciousness and Cognition*, *Psyche*), sociedades profesionales (*Association for the Study of Consciousness –ASSC*) y conferencias anuales dedicadas exclusivamente a su investigación (*Toward a Science of Consciousness*, ASSC).

2. Los Conceptos de Conciencia

Las palabras “consciente” y “conciencia” son términos sombrilla que cubren una amplia variedad de fenómenos mentales. Ambos se usan con una diversidad de significados, y el adjetivo “consciente” es heterogéneo en su aplicación, pues vale para organismos –conciencia de una criatura- y a estados y procesos mentales particulares –conciencia de estados (Rosenthal, 1986; Gennaro, 1995; Carruthers, 2000).

2.1 Conciencia de las criaturas

Un animal, una persona u otro sistema cognitivo se pueden considerar conscientes en varios sentidos diferentes.

Sensibilidad. Se puede ser consciente en el sentido general de ser simplemente una criatura sensible, capaz de sentir y responder a su mundo (Armstrong, 1981). Ser consciente en este sentido puede admitir grados, no está delimitado tajantemente qué clase de capacidades sensoriales son suficientes. ¿Son conscientes los peces en el aspecto relevante? ¿Qué pasa con las gambas o las abejas?

Vigilia. Se podría además exigir que el organismo esté ejerciendo de hecho esa capacidad en vez de simplemente tener la capacidad o disposición de hacerlo. Así, algo podría contar como consciente sólo si está despierto y en un estado de alerta normal. En ese sentido no contaría como conscientes los organismos que están dormidos o en alguno de los niveles profundos de coma. También aquí los límites son borrosos y puede haber casos intermedios. Por ejemplo, ¿se está consciente en el sentido pertinente cuando se está soñando, o hipnotizado o fatigado?

Autoconciencia. Un tercer sentido, más exigente, podría definir a las criaturas conscientes como aquellas que no solo se percatan, sino que se percatan de que se percatan, y así trataríamos la conciencia como una forma de *autoconciencia* (Carruthers, 2000). El requisito de la autoconciencia se podría interpretar de varias maneras, y con ello variaría qué criaturas cuentan como conscientes. Si se entiende que implica autoconciencia conceptual explícita, muchos animales no humanos e incluso los niños pequeños podrían no estar incluidos, pero si solo se exigen formas más rudimentarias e implícitas de autoconciencia, entonces podrían contar como autoconscientes una amplia gama de criaturas no lingüísticas.

¿Cómo es ser uno? El famoso criterio de Thomas Nagel (1974) de “Cómo es ser uno” intenta capturar otra noción quizá más subjetiva de ser un organismo consciente. Según Nagel, un ser es consciente si hay un “cómo” que es ser esa cosa; es decir, hay cómo parece o aparece subjetivamente el mundo desde el punto de vista mental o experiencial de esa criatura. En el ejemplo de Nagel, los murciélagos son conscientes porque cuando un murciélago experimenta su mundo por medio de su sentido de la ecolocalización, hay un “cómo” que es el experimental por parte del murciélago, aunque los humanos, desde nuestro punto de vista humano, no podamos decididamente entender cómo es ese modo de conciencia desde el propio punto de vista del murciélago.

Sujeto de estados conscientes. Una alternativa, quinta, sería definir la noción de organismo consciente en términos de los estados conscientes. Esto es, se podría definir primero que hace de un estado mental un estado mental consciente y después definir a las criaturas conscientes porque tienen esos estados. Entonces nuestro concepto de organismo consciente dependería del tratamiento que diéramos de los estados conscientes (sección 2.2)

Conciencia transitiva. Además de describir a las criaturas como conscientes en estos diversos sentidos, hay otros sentidos relacionados en los que se dice que las criaturas son *conscientes* de unas u otras cosas. A veces se señala esta distinción como distinción entre las nociones *transitiva* e *intransitiva* de la conciencia, la primera de las cuales involucra algún objeto hacia el que está dirigida la conciencia (Rosenthal, 1986).

2.2 Conciencia de estados.

La noción de estado mental consciente tiene también una variedad de significados distintos aunque quizá interrelacionados. He aquí por lo menos seis de las posibilidades principales.

Estados de los que uno se percata. Según una interpretación común, un estado mental consciente es simplemente un estado mental del que uno se percata que está en él (Rosenthal; 1986, 1996). En este sentido los estados conscientes involucran una forma de *meta-mentalidad* o *meta-intencionalidad*, en la medida en que requiere estados mentales acerca de estados mentales. Tener un deseo consciente de una taza de café es tener ese deseo y también percatarse simultánea y directamente de que uno tiene ese deseo. En este sentido los pensamientos y los deseos inconscientes no son sino los que tenemos sin percatarnos de tenerlos, sea que nuestra falta de autoconocimiento resulte de la falta de atención o de causas psicoanalíticas más profundas.

Estados cualitativos. Se podría considerar conscientes a los estados en un sentido al parecer distinto y más *cualitativo*. Esto es, podríamos contar entre los conscientes a un estado sólo si tiene o involucra

propiedades cualitativas o experienciales del tipo al que a veces se denominan “qualia” o “sensaciones puras”. La percepción que uno tiene del Merlot que está bebiendo o del tejido que está examinando cuentan como estados mentales conscientes porque involucran varios qualia sensoriales; por ejemplo, el qualia del sabor en el caso del vino y el qualia del color en la experiencia visual de la tela. Hay un gran desacuerdo sobre la naturaleza de tales qualia (Churchland, 1985; Shoemaker, 1990; Clark, 1993; Chalmers, 1996) e incluso sobre su existencia. Tradicionalmente los qualia se han considerado como rasgos intrínsecos, privados, inefables de la experiencia, pero las teorías actuales de los qualia a veces rechazan algunos de estas características (Dennett, 1990).

Estados fenoménicos. Hay quien a veces se refiere a estos qualia como a propiedades fenoménicas, y al tipo de conciencia correspondiente como conciencia *fenoménica*, pero este término se aplica quizá con más propiedad a la estructura global de la experiencia e incluso mucho más que los meros qualia sensoriales. La estructura fenoménica de la conciencia también abarca gran parte de la organización espacial, temporal y conceptual de nuestra experiencia del mundo y de nosotros mismos como agentes. (Véase la sección 4.3) Por lo tanto, lo mejor es probablemente, por lo menos al principio, distinguir el concepto de conciencia fenoménica del de conciencia cualitativa, aunque sin duda se superponen.

Estados de cómo es ser uno. La conciencia en los dos sentidos anteriores enlaza bien con la noción de criatura consciente de Thomas Nagel (1974), en la medida en que podríamos considerar consciente un estado mental en el sentido de “cómo es ser uno” sencillamente si estar en ese estado “es como” algo. Se podría entender que el criterio de Nagel intenta proporcionar una concepción interna o de primera persona de qué hace que un estado sea fenoménico o cualitativo.

Conciencia de acceso. Los estados pueden ser conscientes en el sentido, al parecer muy diferente, de conciencia de acceso, que tiene más que ver con las relaciones intra-mentales. A este propósito, que un estado sea consciente depende de su disponibilidad para interactuar con otros estados y del acceso que uno tiene a su contenido. En este sentido más funcional, que se corresponde con lo que Ned Block (1995) llama *conciencia de acceso*, que un estado visual sea consciente no es tanto cuestión de que tenga o no un cualitativo “cómo es estar en uno”, sino de si la información visual que porta está disponible para uso y guía del organismo. En cuanto que la información de ese estado esté disponible, con riqueza y flexibilidad, para el organismo que la contiene, contará como un estado consciente en el aspecto relevante, haya o no una sensación cualitativa o fenoménica en el sentido de Nagel.

Conciencia narrativa. Se pueden considerar conscientes a los estados en un sentido narrativo que apela a la noción de “flujo de la conciencia”, considerado como una narración más o menos serial de episodios desde la perspectiva de un yo real o simplemente virtual. La idea sería identificar los estados conscientes de la persona con los que aparecen en este flujo (Dennett, 1991, 1992).

Aunque estas seis nociones de lo que hace consciente a un estado se pueden especificar independientemente, no carecen, desde luego, de enlaces potenciales, ni agotan el ámbito de las posibilidades. Buscando conexiones, se podría argumentar que los estados aparecen en el flujo de la conciencia solo si nos percatamos de ellos, y así podríamos establecer un lazo entre la primera noción meta-mental de estado consciente y el concepto del flujo o la narración. También podríamos conectar el acceso con las nociones cualitativa o fenoménica de estado consciente intentando mostrar que los estados que representan de esas maneras hacen disponibles sus contenidos de las maneras requeridas por la noción de acceso.

Yendo más allá de estas seis posibilidades, podríamos distinguir entre los estados conscientes y los no conscientes apelando a aspectos de sus dinámicas e interacciones intra-mentales, además de las meras relaciones de acceso; por ejemplo, puede que los estados conscientes manifiesten un acerbo más rico de interacciones sensibles al contenido o un grado mayor de orientación flexible hacia objetivos, del tipo asociado con el control autoconsciente del pensamiento. O, por el contrario, podríamos intentar definir los estados conscientes en términos de las criaturas conscientes. Esto es, podríamos intentar explicar qué es ser una criatura consciente o incluso un yo consciente y luego definir la noción de estado consciente en términos de ser un estado de tal criatura o sistema, que sería lo inverso de la última opción considerada más atrás para definir las criaturas conscientes en términos de los estados mentales conscientes.

2.3 La conciencia como entidad.

También el nombre “conciencia” tiene una variedad de significados que en gran medida son paralelos a los del adjetivo “consciente”. Se pueden hacer distinciones entre la conciencia de la criatura y el estado de conciencia, y entre las variedades de cada una de ellas. Nos podemos referir específicamente a la conciencia fenoménica, a la conciencia de acceso, reflexiva o meta-mental, y a la conciencia narrativa, entre otras.

Aquí la conciencia no se trata generalmente como una entidad sustantiva, sino simplemente como la reificación abstracta de cualquier propiedad o aspecto que se atribuya mediante el uso relevante del adjetivo “consciente”. La conciencia de acceso es simplemente la propiedad de tener el tipo adecuado de relaciones de acceso interno, y la conciencia cualitativa es más que la propiedad que se atribuye cuando “consciente” se aplica en el sentido cualitativo a los estados mentales. Si esto nos compromete con el estatuto ontológico de la conciencia en cuanto tal dependerá de si somos platónicos acerca de los universales en general. No nos tiene por que comprometer con la idea de que la conciencia sea una entidad distinta, o no más que el uso de “cuadrado” “rojo” o “amable” nos compromete con la existencia de la *cuadracidad*, rojez o amabilidad como entidades distintas.

Aunque no es lo corriente, se puede, sin embargo, adoptar una posición realista más robusta acerca de la conciencia como componente de la realidad. O sea, podemos concebir la conciencia más al estilo de los campos electromagnéticos que al estilo de la vida.

Desde que fue rechazado el vitalismo, no entendemos la vida *per se* como algo distinto de los seres vivos. Hay cosas vivas, incluyendo organismos, estados, propiedades y o partes de organismos, comunidades y linajes evolutivos de organismos, pero la vida no es otra cosa más, un componente adicional de la realidad, alguna fuerza vital que se añade a los seres vivos. Aplicamos los adjetivos “vivo” y “viviente” correctamente a muchas cosas, y al hacerlo se puede decir que les atribuimos vida, pero sin otro significado ni otra realidad aparte de la involucrada en que son cosas vivas.

Por el contrario, los campos electromagnéticos se consideran partes reales e independientes de nuestro mundo físico. Aunque a veces podamos especificar los valores de estos campos apelando a la conducta de las partículas que están en él, los campos mismos se consideran como constituyentes concretos de la realidad, y no simplemente como abstracciones o conjuntos de relaciones entre partículas.

De manera semejante, podríamos entender que “conciencia” se refiere a un componente o aspecto de la realidad que se manifiesta en los estados y criaturas conscientes, pero es algo más que la nominalización abstracta del adjetivo “consciente” que les aplicamos. Aunque tales opiniones realistas fuertes no son muy comunes hoy día, hay que incluirlas en el espacio lógico de la elección.

Hay, pues, muchos conceptos de conciencia, y tanto “consciente” como “conciencia” se usan de diversas maneras sin que haya un significado privilegiado o canónico. Sin embargo, esto más que un impedimento es riqueza de impedimenta. La conciencia es un rasgo complejo del mundo, y entenderla requerirá una diversidad de herramientas conceptuales para tratar sus numerosos aspectos diferentes. Por lo tanto, la pluralidad conceptual es lo que habría que esperar. Mientras evitemos la confusión teniendo claros los significados, es bueno tener una variedad de conceptos por medio de los cuales acceder a la conciencia y captar toda su rica complejidad. Sin embargo no tenemos por qué presuponer que la pluralidad conceptual implica divergencia referencial. Puede que nuestros múltiples conceptos de conciencia en realidad seleccionen varios aspectos de un único y singular fenómeno mental subyacente. Si esto es así y en qué medida es una cuestión abierta.

3. Los problemas de la conciencia

También la tarea de entender la conciencia es un proyecto diverso. No es solo que haya muchos aspectos de la mente que se puedan considerar conscientes en algún sentido, además, cada uno de ellos, es susceptible de explicación o modelización en varios aspectos. Entender la conciencia involucra una multiplicidad no solo de *explananda* sino también de cuestiones que plantean y el tipo de respuestas que requieren. A riesgo de simplificar, las preguntas pertinentes se pueden reunir bajo tres amplias rúbricas: la cuestión del qué, la del cómo y la del porqué:

· La Cuestión Descriptiva: ¿Qué es la conciencia?, ¿cuáles son sus características principales? Y ¿cuáles son los mejores medios para descubrirlas, descubrirlas y modelarlas?

· La Cuestión Explicativa: ¿Cómo es que existe la conciencia del tipo pertinente? ¿Es un aspecto primitivo de la realidad y, si no, cómo surge (o puede surgir) la conciencia, en el aspecto relevante, a partir de entidades o procesos no conscientes o es causada por ellos?

· La Cuestión Funcional: ¿Por qué existe la conciencia del tipo relevante? ¿Tiene alguna función, y en caso afirmativo, cuál? ¿Actúa causalmente, y en caso afirmativo, qué clase de efectos tiene? ¿Aporta algo a la manera de operar de los sistemas en los que está presente, y en caso afirmativo, por qué y cómo?

Las tres preguntas se centran respectivamente en describir los rasgos de la conciencia, explicar su base o causa subyacentes y explicar su papel o valor. Las divisiones entre las tres son, desde luego, un poco artificiales y, en la práctica, la respuesta que demos a cada una dependerá en parte de lo que digamos de las otras. Por ejemplo, no podemos responder adecuadamente la cuestión del qué y describir los rasgos principales de la conciencia sin abordar la cuestión del porqué de su papel funcional en los sistemas a cuyas operaciones afecta. Tampoco podemos explicar cómo el tipo relevante de conciencia puede surgir a partir de procesos no consciente a menos que tengamos una descripción clara de cuáles son los rasgos que hay que causar o realizar para que cuenten como su producción. A pesar de estas precauciones, la división en tres tipos de cuestiones ofrece una estructura útil para articular el proyecto explicativo general y para evaluar la adecuación de las teorías o modelos particulares de la conciencia.

4. La cuestión descriptiva: ¿Cuáles son los rasgos de la conciencia?

La cuestión del qué nos pide describir y dar un modelo de los rasgos principales de la conciencia, pero qué rasgos son pertinentes dependerá de la clase de conciencia que pretendamos captar. Las propiedades principales de la conciencia de acceso pueden ser muy distintas de las de la conciencia cualitativa o fenoménica, y las de la conciencia reflexiva o de la conciencia narrativa pueden ser diferentes de ambas. Sin embargo, construyendo teorías detalladas para cada tipo, podemos tener la esperanza de hallar lazos importantes entre ellas y quizá descubrir que coinciden en al menos algunos aspectos clave.

4.1 Datos de primera y de tercera persona

El proyecto general descriptivo exigirá una variedad de métodos de investigación (Flanagan, 1992). Aunque podríamos ingenuamente considerar los hechos de la conciencia como demasiado evidentes para exigir métodos sistemáticos de acopio de datos, la tarea epistémica está muy lejos de ser trivial (Husserl, 1913).

El acceso introspectivo de primera persona suministra una fuente rica de conocimientos sobre nuestra vida mental consciente, pero no es suficiente por sí mismo ni es especialmente útil a menos que se use de una manera disciplinada y entrenada. Reunir los datos necesarios sobre la estructura de la experiencia consciente exige una perspectiva rigurosa, dirigida hacia el interior, que es muy distinta de nuestra forma cotidiana de percatarnos de nosotros mismos (Husserl, 1929; Merleau-Ponty, 1945). La observación hábil del tipo necesario requiere adiestramiento, esfuerzo y capacidad para adoptar diferentes perspectivas sobre la experiencia propia.

La necesidad de datos empíricos de tercera persona aportados por observadores externos es quizá más evidente cuando tratamos de los tipos más claramente funcionales de conciencia, tales como la conciencia de acceso, pero se necesitan incluso en el caso de la conciencia cualitativa y fenoménica. Por ejemplo, los estudios de déficits que correlacionan varias áreas neuronales y funcionales de lesiones como anomalías de la experiencia consciente pueden advertirnos de aspectos de la estructura fenoménica que escapan de nuestra conciencia introspectiva normal. Como muestra el estudio de estos casos, en la experiencia pueden venir separadas cosas que parecen invariablemente unificadas o singulares desde nuestro punto de vista normal de primera persona (Sacks, 1985; Shallice, 1988; Farah, 1995).

O, por poner otro ejemplo, los datos de tercera persona pueden ayudarnos a ver cómo nuestras experiencias de actuar y nuestras experiencias de registrar el tiempo se afectan las unas a las otras de maneras que nunca podríamos averiguar por simple introspección (Libet, 1985; Wegner, 2002). Tampoco se limitan los hechos aportados por estos métodos de tercera persona a las causas o bases de la conciencia, sino que a veces afectan a la estructura misma de la conciencia fenoménica. Para reunir los datos necesarios se

necesitarán los métodos de primera persona, de tercera persona y quizá incluso los métodos interactivos de segunda persona (Varela, 1995).

Empleando todas estas fuentes de datos esperamos poder construir modelos descriptivos detallados de los varios tipos de conciencia. Aunque haya rasgos específicos de la mayor importancia que puedan variar entre los diferentes tipos, nuestro proyecto descriptivo general tendrá que abordar al menos los siguientes siete aspectos generales de la conciencia (secciones 4.2-4.7)

4.2 Carácter cualitativo

El *carácter cualitativo* se identifica a veces con las llamadas “sensaciones puras” y se ilustra con la rojez que se experimente cuando se mira a los tomates maduros o al sabor específico que se disfruta cuando se saborea una piña madura (Locke, 1688). El tipo pertinente de carácter cualitativo no se restringe a los estados sensoriales, sino que se entiende que está presente como aspecto de los estados experienciales en general, como los pensamientos o los deseos experimentados (Siewert, 1998).

Puede que a alguien le parezca que la existencia de estas sensaciones marca el umbral de los estados o de las criaturas que son realmente conscientes. Si un organismo siente y responde de maneras apropiadas a su mundo, pero carece de tales qualia, entonces podría contar como consciente, como mucho, en un sentido amplio y menos que literal. O esto dirían quienes toman la conciencia cualitativa, en el sentido de “cómo es ser un”, como filosófica y científicamente central (Nagel, 1974; Chalmers, 1996).

Los problemas relativos a los qualia -¿Puede haber qualia invertidos? (Block, 1980^a, 1980b; Shoemaker, 1981, 1982). ¿Son los qualia epifenómenos? (Jackson, 1982; Chalmers, 1996). ¿Cómo pueden los estados neuronales originar qualia? (Levine, 1983; McGuinn, 1991)- han tenido un gran protagonismo en el pasado reciente. Pero la cuestión del qué suscita un problema de los qualia más básico, a saber: el de dar una descripción clara y articulada de nuestro espacio de los qualia y el estatuto de los qualia específicos dentro de él.

A falta de tal modelo es muy fácil caer en errores factuales o descriptivos. Por ejemplo, las afirmaciones de que es ininteligible la conexión entre el rojo experimentado y cualquier posible sustrato neuronal de esa experiencia tratan a veces el quale relevante de color como una propiedad simple y *sui generis* (Levine, 1983), pero la rojez fenoménica existe en realidad en un espacio de los colores complejo que tiene muchas dimensiones y relaciones de semejanza sistemáticas (Hardin, 1992). Entender el quale específico de color con relación a una estructura relacional más amplia no solo nos da una descripción mejor de su carácter cualitativo, sino que puede suministrar “perchas” de las que colgar lazos inteligibles psico-físicos.

Puede que el color sea la excepción porque de su espacio cualitativo tenemos una comprensión formal específica y bien desarrollada, pero no es la excepción en cuanto a la importancia de tales espacios para comprender las propiedades cualitativas en general (Clark, 1993; Churchland, 1995).

4.3 La estructura fenoménica

La *estructura fenoménica* no se debe confundir con la estructura cualitativa, a pesar de que a veces se intercambia el uso de “qualia” y “propiedades fenoménicas”. La “organización fenoménica” abarca todas las clases de orden y de estructura que se encuentran en el dominio de la experiencia, esto es, en el dominio del mundo tal como se nos *aparece*. Hay, evidentemente, conexiones importantes entre lo fenoménico y lo cualitativo. Ciertamente los qualia se deberían acaso entender como propiedades de los objetos fenoménicos o experimentados, pero lo fenoménico es algo más que la sensación pura. Como Kant (1787), Husserl (1913) y generaciones de fenomenólogos han mostrado, la estructura fenoménica de la experiencia tiene una gran riqueza intencional y no solo involucra las ideas y cualidades sensoriales, sino representaciones complejas del tiempo, el espacio, la causalidad, el cuerpo, el yo, el mundo y la estructura organizada de la realidad vivida en todas sus formas conceptuales y no conceptuales.

Puesto que muchos estados no conscientes tienen también aspectos intencionales y representacionales, quizá lo mejor sea entender que la estructura fenoménica incluye un tipo especial de organización y contenido intencional y representacional, el tipo asociado distintivamente con la conciencia (Siewert, 1998).

Responder a la cuestión de qué exige una cuidadosa descripción del marco representacional, coherente y densamente organizado, en el cual se incardinan las experiencias particulares. Puesto que la mayor parte de esa estructura está solo implícita en la organización de la experiencia, no se puede simplemente “leer”

mediante la introspección. Articular de manera clara e inteligible la estructura del dominio fenoménico es un proceso largo y difícil de inferencia y de modelización (Husserl, 1929). La introspección puede ayudar, pero también hace falta mucha teorización e ingenio.

4.4 La subjetividad

La noción de *subjetividad* es otra que a veces se identifica con los aspectos cualitativos y fenoménicos de la conciencia, pero también en este caso hay buenas razones para reconocer que, al menos en alguna de sus formas, es un rasgo distinto de la conciencia, relacionado con lo cualitativo y lo fenoménico, pero diferente de ambos. En particular, la forma epistémica de la subjetividad tiene que ver con los límites aparentes de la cognoscibilidad o incluso de la inteligibilidad de varios hechos relativos a la experiencia consciente (Nagel, 1974; Van Gulick, 1985; Lycan, 1996).

Según Thomas Nagel (1974) los hechos acerca de cómo es ser un murciélago son subjetivos en el sentido pertinente porque solo se pueden entender completamente desde el punto de vista típico del murciélago. Solo las criaturas capaces de tener o soportar experiencias semejantes pueden entender su “cómo es ser un” en el sentido empático requerido. Los hechos relativos a la experiencia consciente no se pueden entender más que incompletamente desde el punto de vista externo de la tercera persona, como es el asociado con la ciencia física objetiva. Una opinión similar sobre los límites de las teorías de tercera persona es la que parece estar detrás de las afirmaciones sobre lo que la hipotética Mary de Frank Jackson (1982), la supercientífica experta en el color, no podría entender sobre la experiencia del rojo, debido a su propia experiencia visual acromática.

Está abierto a debate si los hechos relativos a la experiencia están limitados epistémicamente (Lycan, 1996), pero la tesis de que entender la conciencia requiere formas especiales de conocer y de acceder desde el punto de vista interno es plausible intuitivamente y tiene una larga historia (Locke, 1688). Así, cualquier respuesta adecuada a la cuestión del qué debe abordar el estatuto epistémico de la conciencia, tanto nuestras capacidades para entenderla como nuestros límites (Papineau, 2002; Chalmers, 1003).

4.5 La organización desde la perspectiva del yo

La estructura perspectivista de la conciencia es un aspecto de su organización fenoménica global, pero tiene importancia suficiente como para merecer un estudio dedicado exclusivamente a ella. En la medida en que la perspectiva clave es la del yo consciente, este rasgo específico se podría llamar *auto-perspectividad*. Las experiencias conscientes no existen como átomos mentales aislados, sino como modos o estados de un yo o sujeto consciente (Descartes, 1644; Searle, 1992, a pesar de Hume, 1739). Una experiencia visual de una esfera azul es siempre cosa de que haya algún yo o sujeto al que le aparece. Un dolor agudo siempre es un dolor sentido o experimentado por un sujeto consciente. El yo no necesita aparecer como un elemento explícito de nuestras experiencias, pero como Kant (1787) advirtió, el “yo pienso” tiene al menos que poder acompañar a cada una de ellas.

El yo se podría entender como el punto de perspectiva desde el cual el mundo de los objetos se presenta a la experiencia (Wittgenstein, 1921). No solo proporciona una perspectiva espacial y temporal para nuestra experiencia del mundo, también una perspectiva de significado e inteligibilidad. La coherencia intencional del dominio experiencial se basa en la interdependencia dual entre el yo y el mundo: el yo como perspectiva desde la cual se conocen los objetos y el mundo como una estructura integrada de objetos y acontecimientos cuyas posibilidades de ser experimentados definen implícitamente la naturaleza y la localización del yo (Kant, 1787; Husserl, 1929).

Es claro que los organismos conscientes difieren en el grado en que constituyen un yo unificado y coherente, y es probable que en consecuencia difieran en la clase o grado de focalización de la perspectiva que incorporan en sus respectivas formas de experiencia (Lorenz, 1977). Puede que la conciencia no requiera un yo distinto o sustancial de tipo cartesiano, pero parece que al menos cierto grado de auto-organización con una perspectiva es esencial para la existencia de algo que cuente como experiencia consciente. Las experiencias parecen tan incapaces de existir sin un yo o sujeto como las olas sin el mar. La cuestión descriptiva, por lo tanto, requiere tratar el aspecto de la experiencia que es la perspectiva del yo y la

organización en un yo de la mente consciente de la cual depende, incluso aunque ese tratamiento considere el yo de una manera relativamente deflacionista y virtual (Dennett, 1991, 1992).

4.6 Unidad

La *unidad* está estrechamente conectada con la perspectiva del yo, pero merece una mención específica como aspecto clave de la organización de la conciencia. Los sistemas conscientes y los estados mentales conscientes implican muchas formas de unidad. Algunas son formas causales de unidad asociadas con la integración de la acción y del control en un foco unificado de agencia. Otras son formas de unidad más representacionales e intencionales que implican la integración de diversos elementos de contenido a varias escalas y niveles de conexión (Cleeremans, 2003).

Algunas de esas integraciones son relativamente locales, como cuando las diversas características detectadas en modalidades sensoriales singulares se combinan en representaciones de objetos externos que poseen esos rasgos, por ejemplo, cuando tenemos una experiencia visual consciente de una lata roja de sopa que pasa por encima de una servilleta verde a rayas (Tries y Gelade, 1980).

Hay otras formas de unidad intencional que abarcan un ámbito más amplio de contenidos. El contenido de la experiencia presente de la habitación en la que estamos sentados depende en parte de su localización en una estructura mucho más amplia asociada con que nos damos cuenta de nuestra existencia como observadores en el tiempo, dentro de un mundo de objetos conectados espacialmente y de existencia independiente (Kant, 1787; Husserl, 1913). El contenido de la experiencia individual solo es posible porque reside en una estructura más amplia de representaciones.

4.7 Intencionalidad y transparencia

Normalmente se entiende que los estados mentales conscientes tienen un aspecto representacional o intencional, en la medida en que versan acerca de cosas, se refieren a cosas o tienen condiciones de satisfacción. Una experiencia visual consciente nuestra *representa correctamente* el mundo si hay lilas en un florero sobre la mesa (a pesar de Travis, 2004), un recuerdo consciente lo es *del* ataque al World Trade Center, y un deseo consciente es *de* un vaso de agua fría. Sin embargo, los estados no conscientes también pueden tener intencionalidad de esta manera, y es importante entender las maneras en que los aspectos representacionales de los estados conscientes se parecen y se diferencian de los de los estados no conscientes (Carruthers, 2000). Searle (1990) propone una opinión contraria según la cual solo los estados conscientes y las disposiciones a tener estados conscientes pueden ser verdaderamente intencionales, pero la mayoría de los teorizadores consideran que la intencionalidad se extiende al dominio inconsciente.

Una dimensión importante potencialmente en la que difieren es la llamada *transparencia*, que es un rasgo importante de la conciencia en dos sentidos metafóricos interrelacionados, cada uno de los cuales tiene aspectos intencionales, experienciales y funcionales.

A menudo se dice que la experiencia perceptual consciente es transparente o, como dice G.E. Moore (1922), “diáfana”. Transparente porque “miramos a través” de nuestra experiencia sensorial en cuanto que parece que nos percatamos directamente de los objetos y acontecimientos externos presentes ante nosotros, en vez de percatarnos de las propiedades de la experiencia por medio de las cuales ella nos presenta o representa esos objetos. Cuando miro los prados azotados por el viento me percató de la hierba verde que ondula y no de una propiedad de mi experiencia visual. Moore mismo creía que podemos percatarnos de estas propiedades con esfuerzo y redirigiendo la atención, aunque algunos partidarios contemporáneos de la transparencia lo niegan (Harman, 1990; Tye, 1992).

Los pensamientos y experiencias conscientes son transparentes además en un sentido semántico, porque sus significados nos parecen conocidos inmediatamente en el mismo acto de pensar en ellos (Van Gulick, 1992). En este sentido se podría decir que “pensamos a través” de ellos en lo que significan o representan. La transparencia en este sentido semántico puede corresponder, por lo menos en parte, con lo que John Searle denomina la “intencionalidad intrínseca” de la conciencia (Searle, 1992).

Nuestros estados mentales conscientes parecen poseer sus significados intrínsecamente, o procedentes de dentro nada más que por lo que son en sí mismos, contra lo que sostienen muchas teorías externalistas del

contenido mental que fundan el contenido en las relaciones entre los portadores de intencionalidad y sus objetos semánticos o referenciales.

La idea de que el contenido mental está determinado intrínsecamente y es auto-evidente desde un punto de vista interno se defiende a veces apelando a nuestras intuiciones sobre “los cerebros en bañeras”, que nos inducen a pensar que los estados mentales conscientes de un cerebro en una bañera mantendrían todo su contenido intencional normal a pesar de perderse los ordinarios lazos causales e informacionales con el mundo (Horgan y Tienson, 2002). Sigue en curso una controversia sobre estos casos y sobre las doctrinas rivales internalistas (Searle, 1992) y externalistas (Dretske, 1995) sobre la intencionalidad consciente.

Aunque la transparencia semántica y la intencionalidad intrínseca tienen algunas afinidades no se han de equiparar llanamente, puesto que es posible acomodar la primera de las nociones con una interpretación más externalista del contenido y el significado. Las transparencias semántica y sensorial aluden desde luego a los aspectos representacionales o intencionales de la conciencia, pero también son aspectos experienciales de nuestra vida consciente. Son parte del “cómo es ser uno” o cómo se siente fenomenológicamente el ser consciente. Ambas tienen aspectos funcionales, en cuanto que las experiencias conscientes interactúan abundantemente entre sí por razón de su contenido, y esas interacciones manifiestan nuestra comprensión transparente de su contenido.

4.8 Flujo dinámico

La dinámica de la conciencia se manifiesta en el orden coherente de su proceso de flujo y transformación, que William James (1800) llamaba “la corriente de la conciencia”. Algunas sucesiones temporales de la experiencia son generadas por factores puramente internos, como cuando pensamos en un rompecabezas, y otras dependen en parte de causas externas, como cuando cogemos una pelota al vuelo, pero incluso estas sucesiones vienen conformadas en gran parte por cómo la conciencia se transforma a sí misma.

Ya sea en parte como respuesta a las influencias externas o venga por entero del interior, cada momento de la sucesión de la experiencia se desarrolla coherentemente a partir de los que le precedieron, restringido y habilitado por la estructura general de enlaces y límites incorporados en su organización previa subyacente (Husserl, 1913). A este propósito la conciencia es un sistema autopoietico; es decir, un sistema que se crea y se organiza a sí mismo (Varela y Maturana, 1980).

Como agente mental consciente puedo hacer cosas como revisar mi habitación, revisar una imagen mental de ella, repasar en la memoria los platos de una comida reciente en un restaurante y muchos de los sabores y aromas, razonar la solución de un problema complejo o planear mis compras y ejecutar el plan cuando llego al mercado. Todas estas son actividades rutinarias y comunes, pero cada una de ellas conlleva la generación dirigida de experiencias de maneras que manifiestan una comprensión práctica implícita de sus propiedades intencionales y de sus contenidos interconectados (Van Gulick, 2000).

La conciencia es un proceso dinámico y, por lo tanto, una respuesta descriptiva adecuada a la pregunta del qué tiene que ocuparse de algo más que de sus propiedades estáticas o momentáneas. En particular, tiene que dar cuenta de la dinámica temporal de la conciencia y de las maneras como su flujo de autotransformaciones refleja su coherencia intencional y la comprensión semántica incorporada en los controles organizadas por medio de los cuales las mentes conscientes se rehacen a sí mismas continuamente como sistemas autopoieticos comprometidos con sus mundos.

Un tratamiento descriptivo adecuado de la conciencia no podría limitarse a estos siete rasgos, pero tenerlos claros nos haría avanzar mucho hacia la respuesta a “qué es la conciencia”.

5. La cuestión explicativa: ¿Cómo es que existe la conciencia?

La cuestión del *cómo* se centra en la explicación más que en la descripción. Nos pide explicar el estatuto básico de la conciencia y su lugar en la naturaleza. ¿Es un rasgo fundamental de la naturaleza por derecho propio o depende su existencia de otros elementos no conscientes, sean físicos, biológicos, neuronales o computacionales? Y en el segundo caso, ¿podemos explicar o entender cómo los elementos relevantes no conscientes pueden causar o realizar la conciencia?

En pocas palabras, ¿podemos explicar cómo se hace algo consciente a partir de cosas que no son conscientes?

5.1 La diversidad de proyectos explicativos

La cuestión del cómo no es una cuestión única, sino una familia general de cuestiones más específicas (Van Gulick, 1995). Todos se ocupan de la posibilidad de explicar uno u otro aspecto de la conciencia, pero varían en sus *explananda* particulares, las restricciones sobre los *explananda* y los criterios de cuál es explicación acertada. Por ejemplo, podríamos preguntar si podemos explicar computacionalmente la conciencia de acceso imitando en un modelo computacional las relaciones de acceso requeridas. O podríamos interesarnos en si las propiedades fenoménicas y cualitativas de la mente de una criatura consciente se pueden deducir *a priori* a partir de una descripción de las propiedades neuronales de sus procesos cerebrales. Ambas son versiones de la cuestión del cómo, pero preguntan por la viabilidad de proyectos explicativos muy diferentes y así pueden tener contestaciones distintas (Lycan, 1996). Sería poco práctico, si no imposible, catalogar todas las versiones posibles de la cuestión del cómo, pero se pueden enumerar algunas de las principales posibilidades.

Explananda. Los *explananda* posibles incluirían los varios géneros de conciencia de estados y de criaturas diferenciados más arriba, así como los siete rasgos de la conciencia indicados en respuesta a la cuestión del qué. Esos dos tipos de *explananda* de superponen e intersecan. Podríamos, por ejemplo, intentar explicar el aspecto dinámico de la conciencia fenoménica o de la de acceso. O podríamos intentar explicar la subjetividad de la conciencia cualitativa o de la meta-metal. No todos los géneros de conciencia poseen todas las propiedades, pero todas ellas lo son de algunos de ellos. Cómo se explica una propiedad dada en relación con uno u otro tipo de conciencia puede que no tenga que corresponda con lo que hace falta para explicarla en relación con otro.

Explanans. El dominio de los *explanans* posibles también es diverso. En su forma quizá más amplia, la cuestión del cómo pregunta cómo la conciencia del tipo pertinente podría estar causada o realizada por elementos no conscientes, pero podemos generar otras muchas cuestiones más específicas restringiendo el ámbito de los *explanans* pertinentes. Podríamos buscar la explicación de cómo un rasgo determinado de la conciencia es causado o realizado por procesos *neuronales* o estructuras *biológicas* o mecanismos *físicos* o relaciones *funcionales* o *teleológicas* o por la organización *computacional* o incluso por estados *mentales no conscientes*. La viabilidad del éxito explicativo también variará. En general cuanto más limitado y elemental el alcance del *explanans*, más difícil será el problema de explicar cómo puede bastar para producir la conciencia (Van Gulick, 1995).

Criterios de explicación. El tercer parámetro claro es cómo definimos el criterio de tener una explicación feliz. Podríamos exigir que el *explanandum* se dedujera *a priori* del *explanans*, aunque es discutible que esto sea un criterio necesario o suficiente para explicar la conciencia (Jackson, 1993). Que sea suficiente dependerá en parte del carácter de las premisas a partir de las cuales procede la deducción. Como cuestión de lógica, necesitaremos algunos principios puente que conecten las proposiciones o las oraciones acerca de la conciencia con las que no la mencionan. Si nuestras premisas se refieren a hechos físicos o neuronales, entonces necesitaremos algunos principios puente o enlaces que conecten estos hechos con los hechos acerca de la conciencia (Kim, 1998). Los enlaces simples, sean nómicos o solo correlaciones bien confirmadas, podrían ofrecer un puente lógicamente suficiente para inferir conclusiones acerca de la conciencia. Pero probablemente no nos permitirían ver cómo o por qué se dan esas conexiones y, por lo tanto, no alcanzarían a explicar del todo cómo existe la conciencia (Levine, 1983, 1993; McGinn, 1991).

Legítimamente podríamos pedir más, en particular una explicación que hiciera inteligible por qué se dan esas conexiones y, quizá, por qué no pueden dejar de darse. A veces se apela a un modelo corriente, en dos fases, para explicar macro-propiedades por medio de micro-sustratos. En un primer paso se analiza la macropropiedad en términos de condiciones funcionales, y luego, en un segundo paso, se muestra que las micro-estructuras que obedecen leyes de su propio nivel son suficientes nómicamente para garantizar la satisfacción de las condiciones funcionales relevantes (Armstrong, 1968; Lewis, 1972).

Las micro-propiedades de las colecciones de moléculas de H₂O a 20°C bastan para satisfacer las condiciones para el estado líquido del agua que componen. Además el modelo hace inteligible cómo las

micro-propiedades producen el estado líquido. Podría parecer que una explicación satisfactoria de cómo se produce la conciencia exige un proceso similar de dos fases. Sin este incluso la deducibilidad *a priori* podría parecer explicativamente escaso. Aunque la necesidad de ese proceso sigue siendo controvertida (Block y Stalnaker, 1999; Chalmers y Jackson, 2001).

5.2 El abismo explicativo

Nuestra incapacidad actual para ofrecer un enlace adecuadamente explicativo se describe a veces, siguiendo a Joseph Levin (1983), como un *abismo explicativo*, y como una indicación de lo incompleto de nuestra comprensión de cómo la conciencia puede depender de un sustrato no consciente, en especial de una estructura física. Esta tesis básica del abismo admite muchas variedades en cuanto a su generalidad y, por lo tanto, en cuanto a su fuerza.

En su forma quizá más débil afirma que nuestras capacidades explicativas *actuales* tienen un límite *práctico*; dados nuestros modelos y nuestras teorías actuales no podemos articular un enlace inteligible. Una versión más fuerte hace una afirmación de *principio* sobre nuestras capacidades humanas y afirma que dados los límites cognitivos humanos nunca podremos cruzar ese abismo. Para nosotros, o criaturas cognitivamente semejantes, ha de ser para siempre un misterio (McGinn, 1991). Colin McGinn (1995) ha argumentado que dada la naturaleza inherentemente especial de nuestros conceptos perceptuales y de los conceptos científicos que derivamos de ellos, no somos adecuados conceptualmente para entender la naturaleza del enlace psico-físico. Los hechos que tienen que ver con enlace nos están cerrados conceptualmente, igual que los hechos sobre la multiplicación y las raíces cuadradas lo están para los armadillos. No caen dentro de nuestros repertorios conceptuales u cognitivos. Una versión todavía más fuerte del abismo elimina la restricción a nuestra dotación cognitiva y niega que en *principio* pueda el abismo ser cruzado por *ningún agente cognitivo*.

Quienes afirman que hay un abismo están en desacuerdo sobre cuáles son las conclusiones metafísicas, si hay alguna, que se siguen de nuestros supuestos límites epistémicos. El mismo Levine ha sido reacio a inferir conclusiones ontológicas anti-fisicalistas (Levine, 1993, 2001). Por otro lado, algunos neodualistas han intentado usar la existencia del abismo para refutar el fisicalismo (Foster, 1996; Chalmers, 1996). Cuanto más fuerte sea nuestra premisa epistemológica mayor será la esperanza de obtener una conclusión metafísica. Así que no es de extrañar que las conclusiones dualistas se apoyen a menudo apelando a la supuesta imposibilidad de *principio* de cerrar el abismo.

Si pensamos por razones *a priori* que no hay manera posible de explicar la conciencia a partir de lo físico, no es dar un gran paso concluir que en realidad no surge de lo físico (Chalmers, 1996). Sin embargo, la fuerza misma de tal posición epistemológica es difícil de asumir sin previamente haberse basado en el resultado metafísico en cuestión. Así que quienes quieran usar una tesis fuerte del “abismo en principio” para refutar el fisicalismo, tiene que encontrar fundamentos independientes para sustentarla. Algunos han apelado en su apoyo a argumentos sobre la concebibilidad, como la supuesta concebibilidad de zombies idénticos molecularmente a seres humanos conscientes pero desprovistos de toda conciencia fenoménica (Campbell, 1970; Kirk, 1974; Chalmers, 1996). Otros argumentos aducen el supuesto carácter no funcional de la conciencia y, con ello, su resistencia al método científico que ordinariamente se emplea para explicar las propiedades complejas (por ejemplo, la dominancia genética) en términos de condiciones funcionales realizadas físicamente (Block, 1980^a; Chalmers, 1996). Esos argumentos eluden la circularidad en la cuestión anti-fisicalista, pero se apoyan a su vez en afirmaciones e intuiciones controvertidas y no del todo independientes de las opiniones básicas que el proponente tiene acerca del fisicalismo. Sobre este tema siguen activas las discusiones.

Nuestra actual incapacidad para ver alguna manera de pasar el abismo puede que ejerza alguna influencia sobre nuestras intuiciones, pero puede que solo refleje los límites de nuestras teorías actuales más que una barrera en principio insorteable (Dennett, 1991). Más aun, algunos fisicalistas han argumentado que son de esperar los abismos explicativos, e incluso los implican, algunas versiones plausibles del fisicalismo ontológico: las que tratan a los agentes humanos como sistemas cognitivos realizados físicamente, pero con límites inherentes que derivan de su origen evolutivo y de su modo de comprensión situado contextualmente (Van Gulick, 1985, 2003; McGinn, 1991; Papineau, 1995, 2002). Según esta postura, en vez de refutar el

fisicalismo, puede que la existencia del abismo explicativo lo confirme. Las discusiones y los desacuerdos sobre estos temas continúan.

5.3 Explicación reductiva y no reductiva

Como ha mostrado la necesidad de una conexión inteligible, no está claro que la deducibilidad a priori sea de por sí suficiente para tener éxito en la explicación (Kim, 1980), y tampoco es claramente necesaria. Puede que una relación lógica más débil baste en muchos contextos explicativos. A veces podemos contar lo bastante como para satisfacernos acerca de la historia de cómo los hechos de un tipo dependen de los de otro, aunque no podamos estrictamente deducir los hechos del primer género de los del segundo.

La deducción inter-teórica estricta la tomó la doctrina empirista lógica de la unidad de la ciencia como la norma de reducción (Putnam y Oppenheim, 1958), pero en décadas más recientes ha ganado favor una imagen no reductiva y más débil de las relaciones entre las ciencias. En particular, los materialistas no reduccionistas han defendido la llamada “autonomía de las ciencias especiales” (Fodor, 1974) y la idea de que entender el mundo natural nos exige emplean una variedad de sistemas conceptuales y representacionales que acaso no sean estrictamente traducibles entre sí ni susceptibles de tener las correspondencias estrictas que exigían el antiguo paradigma deductivo de las relaciones entre distintos niveles (Putnam, 1975).

A veces se cita como ejemplo la economía (Fodor, 1974; Searle, 1992). Puede que los hechos económicos sean realizados por procesos físicos subyacentes, pero nadie pide en serio que podamos deducir los hechos económicos relevantes a partir de descripciones detalladas de sus bases físicas subyacentes, ni que podamos poner en estrecha correspondencia los conceptos y el vocabulario de la economía con el de las ciencias físicas.

A pesar de todo, no se interpreta esa incapacidad deductiva como causa de defectos ontológicos; no existe el problema “dinero-materia”. Lo único que hace falta es una comprensión general y menos que deductiva de cómo las propiedades y relaciones económicas podrían sustentarse sobre las físicas. Así, podríamos optar por un criterio semejante para interpretar la cuestión del cómo y para qué cuenta como una explicación de que la conciencia podría ser causada o realizada por elementos no conscientes. Sin embargo, algunos críticos, como Kim (1987), han objetado la coherencia de cualquier tesis que pretenda ser a la vez no reduccionista y fisicalista, aunque a su vez han replicado los defensores de esas tesis (Van Gulick, 1993).

Otros han argumentado que la conciencia se resiste especialmente a una explicación en términos físicos debido a las diferencias inherentes que hay entre nuestros modos de comprensión subjetivos y objetivos. Es famoso que Thomas Nagel (1974) argumentó que hay límites insoslayables para nuestra capacidad de entender la fenomenología de la experiencia del murciélago, y esos límites los pone nuestra incapacidad de adoptar empáticamente un perspectiva experiencial como la que caracteriza a la experiencia auditiva que el murciélago tiene de su mundo gracias a la eco-localización. Dada nuestra incapacidad para sufrir semejante experiencia, podemos como mucho tener una comprensión parcial de la naturaleza de tal experiencia. Por más conocimientos que acopiemos desde la perspectiva de tercera persona de las ciencias naturales, no bastará, se supone, para permitirnos entender lo que el murciélago entiende de su propia experiencia a partir de su punto de vista subjetivo de primera persona.

5.4 Expectativas sobre el éxito explicativo

Así, pues, la cuestión del cómo se subdivide en una familia de cuestiones más específicas, dependiendo del tipo específico o del rasgo de la conciencia que intentemos explicar, las restricciones específicas que pongamos al alcance del *explanans* y el criterio que usemos para definir el éxito explicativo. Algunas de las variantes que resultan parecen más fácil de responder que otras. Parece verosímil tener éxito con algunos de los llamados “problemas fáciles” de la conciencia, tal como explicar la dinámica de la conciencia de acceso en términos de la organización funcional o computacional del cerebro (Baars, 1988). Otros pueden parecer menos tratables, en especial el llamado “problema difícil” (Chalmers, 1995) que es poco más o menos el de dar una explicación inteligible que nos permita ver de manera intuitivamente satisfactoria cómo la conciencia fenoménica o “cómo es ser uno” puede originarse a partir de los procesos físicos o neuronales del cerebro.

Las respuestas positivas a algunas versiones de la cuestión del cómo parecerse hallarse a mano, o casi, pero otras se muestran profundamente desconcertantes. Tampoco debemos suponer que cada versión tenga una respuesta positiva. Si el dualismo es correcto, entonces la conciencia, en al menos algunos de sus tipos, puede ser básica y fundamental. En ese caso no podremos explicar cómo surge de elementos no conscientes, sencillamente porque no lo hace.

Las expectativas que tengamos sobre la explicación de la conciencia dependerán normalmente de nuestra propia perspectiva. Los fisicalistas optimistas probablemente verán las lagunas explicativas como meros reflejos de los estados iniciales de la investigación y estarán seguros de que se llenarán en un futuro no muy lejanos (Dennett, 1991; Searle, 1992; P.M. Churchland, 1995). Para los dualistas estas fronteras significarán la bancarrota del programa fisicalista y la necesidad de reconocer que la conciencia es un constituyente fundamental de la realidad por derecho propio (Robinson, 1982; Foster, 1989, Foster, 1989, 1996; Chalmers, 1996). Lo que uno ve depende en parte de dónde está, y el proyecto de explicar la conciencia estará siempre acompañado por el debate acerca de su estatuto y de sus expectativas de éxito.

6. La cuestión funcional: ¿Por qué existe la conciencia?

La cuestión funcional o del porqué pregunta por el valor o papel de la conciencia y, con ello, por lo tanto, indirectamente acerca de su origen. ¿Tiene una función y, en caso afirmativo, cuál? ¿Aporta algo a la actuación de los sistemas en que está presente y, en caso afirmativo, por qué y cómo? Si la conciencia es un rasgo complejo de los sistemas biológicos, entonces su valor adaptativo verosímilmente ha de ser relevante para explicar su origen evolutivo, aunque, desde luego, su función actual, si tiene alguna, no tiene por qué ser la misma que tuviera cuando se originó. Las funciones adaptativas cambian a menudo en el tiempo biológico. Las cuestiones sobre el valor de la conciencia tienen también una dimensión *moral* en dos direcciones, por lo menos. Nos inclinamos a considerar que el estatuto moral de un organismo está determinado, por lo menos en parte, por el carácter de su conciencia y el grado y en que es consciente, y los estados consciente, en especial los afectivos, tales como los placeres y dolores, desempeñan un gran papel en muchas de las doctrinas del valor que subyacen a la teoría moral (Singer, 1975).

Igual que en el caso de las cuestiones del qué y del cómo, la cuestión del porqué plantea un problema general que se subdivide en una diversidad de indagaciones más específicas. Mientras los varios géneros de conciencia, esto es, de acceso, fenoménica, meta-mental, sean distintos y separables –lo cual sigue siendo una cuestión abierta- también diferirán en sus papeles y valores específicos. Por lo tanto, bien puede ser que la cuestión del porqué no tenga una respuesta única ni uniforme.

6.1 El estatuto causal de la conciencia

Quizá el tema más básico que propone cualquier versión de la cuestión del porqué es si la conciencia tiene o no el tipo pertinente de impacto causal. Si no tiene efectos y no aporta nada causalmente, entonces parece que sería incapaz de desempeñar ningún papel significativo en los sistemas u organismos en los que está presente, y eso cortaría de raíz muchas de las investigaciones sobre su posible valor. Tampoco se puede desdeñar la amenaza de la irrelevancia epifenomenológica como algo imposible, puesto que de al menos algunas formas de conciencia se ha dicho recientemente que carecen de estatuto causal. Estas preocupaciones se han suscitado especialmente a propósito de los qualia y de la conciencia cualitativa (Huxley, 1874; Jackson, 1982; Chalmers, 1996), pero también se han alzado objeciones contra el estatuto causal de otros tipos, incluyendo la conciencia meta-mental (Velmans, 1991).

A favor de estas posiciones se han ofrecido argumentos metafísicos y empíricos. Entre los primeros están los que apelan a la concebibilidad y la posibilidad lógica de los zombies, esto es, seres cuya conducta, organización funcional y estructura física, hasta el nivel molecular son idénticas a las de los agentes humanos normales, pero que carecen de qualia o de conciencia cualitativa. Algunos (Kirk, 1970; Chalmers, 1996) aseguran que esos seres son posibles en mundos que comparten todas nuestras leyes físicas, pero otros lo niegan (Dennett, 1991, Levine, 2001). Si son posibles en tales mundos, entonces parece que se sigue que incluso en nuestro mundo, los qualia no afectan al curso de los acontecimientos físicos, incluyendo los que constituyen nuestras conductas humanas. Si esos acontecimientos se despliegan de la misma manera estén o no presentes los qualia, entonces parece que los qualia serían inertes o epifenoménicos, por lo menos con

respecto a los acontecimientos del mundo físico. Sin embargo, tales argumentos y las intuiciones sobre los zombies de las cuales dependen y su plausibilidad siguen bajo disputa (Searle, 1992; Yablo, 1998; Balog, 1999).

Se han elevado argumentos de tipo mucho más empírico contra el estatuto causal de la conciencia meta-mental, por lo menos en la medida en que su presencia se pueda medir por nuestra capacidad de informar de nuestro propio estado mental. Se afirma que hay indicios científicos que muestran que la conciencia de este tipo no es necesaria para ningún tipo de capacidad mental ni aparece lo bastante pronto como para actuar como causa de los actos o procesos que típicamente consideramos como sus efectos (Velmans, 1991). Según quienes proponen estos argumentos, los tipos de capacidades mentales de los que generalmente pensamos que necesitan de la conciencia se pueden realizar inconscientemente, en ausencia de la supuestamente necesaria percatación de uno mismo.

Es más, incluso cuando la auto-percatación está presente, se afirma que ocurre demasiado tarde como para ser la causa de las acciones relevantes, en vez de una resultado de ellas o, como mucho, un efecto conjunto de una causa anterior compartida (Libet, 1985). La auto-percatación o conciencia meta-mental resulta ser, según estos argumentos un post-efecto psicológico en vez de una causa iniciadora, más parecida a una “salida de impresora” o a lo que aparece en la pantalla del ordenador que a las operaciones reales del procesador que producen la respuesta del ordenador y lo que sale en pantalla.

También estos argumentos son controvertidos, y tanto los supuestos datos como su interpretación son objeto de vivo desacuerdo (véase Flanagan, 1992, y los comentarios que acompañan a Velmans, 1991). Aunque los argumentos empíricos, al igual que el de los zombies, exigen que consideremos seriamente si algunas formas de conciencia pueden no llegar a ser causalmente menos efectivas de lo que generalmente suponemos, muchos teóricos no temen que los datos empíricos sean una amenaza verdadera al estatuto causal de la conciencia.

Si los epifenomenalistas se equivocan y la conciencia, en sus varias formas, es verdaderamente causal, ¿qué clases de efectos tiene y qué aporta? ¿En qué se diferencian los procesos mentales que involucran el tipo pertinente de conciencia de los que no? ¿Qué funciones podría desempeñar la conciencia? Las seis secciones siguientes (6.2-6.7) examinan algunas de las respuestas más comunes. Aunque las diversas funciones se superponen algo, cada una es distinta, y se diferencian también por los tipos de conciencia con los que es más adecuado conectarlas.

6.2 El control flexible

Aumento de la flexibilidad y de la finura del control. Los procesos mentales conscientes parecen suministrar formas de control muy flexibles y adaptativas. Aunque los procesos inconscientes automáticos pueden ser a veces extremadamente eficientes y rápidos, operan típicamente de maneras más fijas y predeterminadas que los que entrañan autoconciencia (Anderson, 1983). La percatación consciente es, pues, de la mayor importancia cuando nos enfrentamos a situaciones nuevas y a problemas o demandas previamente no afrontados (Penfield, 1975; Armstrong, 1981).

Las explicaciones más comunes de la adquisición de habilidades destacan la importancia de la percatación consciente durante la fase inicial del aprendizaje, que da paso gradualmente a procesos más automáticos del tipo de los requieren poca atención o supervisión consciente (Schneider y Shiffrin, 1977). El procesamiento consciente permite la construcción o compilación de rutinas específicamente montadas a partir de unidades elementales y con el control deliberado de su ejecución.

Es conocido que hay una negociación entre la flexibilidad y el control; los procesos conscientes controlados adquieren su versatilidad al precio de ser lentos y penosas, al contrario de la fluida rapidez de las operaciones mentales automáticas inconscientes (Anderson, 1983). Los incrementos relevantes en la flexibilidad parece que están estrechamente relacionados con las formas de conciencia meta-mentales o de orden superior, puesto que la capacidad realzada de controlar los procesos depende de la mayor auto-percatación. Sin embargo la flexibilidad y las formas refinadas de control se pueden también asociar con los tipos de conciencia fenoménico y de acceso.

6.3 La coordinación social

Capacidad realizada de coordinación social. La conciencia de tipo meta-mental bien puede involucrar no solo un incremento de la auto-percatación, sino una comprensión mejorada de los estados mentales de otras criaturas poseedoras de mente, en especial los de otros miembros del grupo social propio (Humphreys, 1982). Las criaturas que poseen la conciencia meta-mental relevante no solo tienen creencias, motivos, percepciones e intenciones, sino que entienden lo que es tener estos estados y se dan cuenta de ellas mismas, y otras también, tienen estos estados.

Esta mejora del mutuo conocimiento de las mentes permite a los organismos interactuar, cooperar y comunicarse de maneras más avanzadas y adaptativas. Aunque la conciencia meta-mental es la que de modo más evidente está conectada con un papel coordinador social, la conciencia narrativa asociada al flujo de la conciencia también es relevante, puesto que entraña la aplicación a uno mismo de la capacidad de interpretación que deriva en parte de su aplicación social (Ryle, 1949; Dennett, 1978, 1992).

6.4 Representación integrada

Representaciones de la realidad más unificadas y más densamente integradas. La experiencia consciente nos presenta un mundo de objetos que tienen una existencia independiente en el espacio y en el tiempo. Esos objetos por lo común se nos presentan en versiones multi-modales que involucran la integración de la información procedente de varios canales sensoriales y de conocimientos y recuerdos de fondo. La experiencia consciente no solo nos presenta propiedades o rasgos aislados, sino objetos y acontecimientos situados en un mundo independiente, y lo hace incorporación en su organización y dinámica experienciales la densa red de relaciones y conexiones que en su conjunto constituyen la estructura significativa de un mundo de objetos (Kant, 1787; Husserl, 1913; Campbell, 1997).

Desde luego, no toda la información sensorial ha de ser experimentada para que tenga un efecto adaptativo sobre la conducta. Se pueden encontrar enlaces adaptativos senso-motores no experienciales en los organismos simples y en algunos de los procesos más directos y reflexivos de los organismos superiores. Pero cuando existe la experiencia, proporciona una representación más unificada e integrada de la realidad, la cual suministra vías más abiertas de respuesta (Lorenz, 1977). Consideremos, por ejemplo, la representación del espacio en un organismo cuyos canales sensoriales de entrada está ligados simplemente al movimiento o a la orientación de unos pocos mecanismos fijos, tales como los de ingestión o de captura de presas, y comparémosla con la de un organismo capaz de usar la información especial para una navegación flexible en su entorno y para cualesquiera metas o fines en los que el espacio cuente, como cuanto una persona supervisa su oficina o su cocina (Gallistel, 1990).

Las representaciones de este tipo son posibilitadas normalmente por el modo integrado de presentación asociado con la experiencia consciente. La unidad del espacio experimentado no es más que un ejemplo del tipo de integración asociada con nuestra percatación consciente de un mundo objetivo.

El papel o valor integrador es el que más evidentemente se asocia con el tipo fenoménico de conciencia, pero en la medida en que los qualia desempeñen un papel en la representación de los objetos ante la experiencia, este papel unificador está también conectado con la conciencia en los sentidos cualitativo y de “cómo es ser uno”. Además está ligado íntimamente con la transparencia de la conciencia descrita en la respuesta a la cuestión del qué, en especial con la transparencia semántica (Van Gulick, 1993).

6.5 Acceso informacional

Un acceso informacional más global, La información que portan los estados mentales conscientes es normalmente asequible para una diversidad de subsistemas mentales y para aplicarla a una amplia gama de situaciones y acciones posibles (Baars, 1988). Lo más corriente es que la información no consciente esté encapsulada en módulos mentales particulares y solamente esté disponible para las aplicaciones conectadas directamente con la operación de esos subsistemas (Fodor, 1983). Hacer consciente la información amplía por lo general su esfera de influencia y los modos en que se puede usar para guiar o conformar adaptativamente la conducta interna o la externa. Que un estado sea consciente puede, en parte, ser asunto de lo que Dennett llama la “popularidad cerebral”, esto es, su capacidad para impactar en virtud de su contenido sobre otros estados mentales.

Esta función particular está conectada directamente, y por definición, con el concepto de conciencia de acceso (Block, 1995), pero también la conciencia meta-mental y las conciencias fenoménica y cualitativa parecen ligadas a semejantes mejoras de la disponibilidad de la información (Armstrong, 1981; Tye, 1985).

6.6 La libertad de la voluntad

Incremento de la libertad de elección o de la libre voluntad. El tema de la libre voluntad sigue siendo un problema filosófico perenne, no solo si la hay o no, sino en qué podría o debería consistir (Dennett, 1984; van Inwagen, 1983; Hasker, 1999; Wegner, 2002). Es posible que la noción de libre voluntad sea demasiado oscura o discutible para proyectar luz sobre el papel de la conciencia, pero existe una intuición tradicional de que las dos estén profundamente conectadas.

Se ha pensado que la conciencia abre un reino de posibilidades, una esfera de alternativas dentro de la cual el yo consciente puede elegir o actuar libremente. Por lo menos parece que la conciencia es un condición necesaria para que exista esa libertad o autodeterminación (Hasker, 1999).

¿Cómo podríamos hacer una elección libre mientras permanecemos en el reino de lo inconsciente? ¿Cómo podemos determinar nuestra voluntad sin ser conscientes de ello y de las alternativas que tenemos para darle forma?

La libertad para elegir las acciones propias y la capacidad para determinar nuestro propio carácter y su desarrollo futuro puede admitir muchos grados y variaciones interesantes y no ser una cuestión de todo o nada, y puede que haya diversas formas y niveles de conciencia que se correlacionen con los grados o tipos correspondientes de libertad o de autodeterminación (Dennett, 1984, 2003). Puesto que destaca aquí la auto-percatación, parece que la forma meta-mental de la conciencia es la más relacionada con la libertad, pero también parece que puede haber conexiones con la mayoría de los otros tipos de conciencia.

6.7 La motivación interna.

Estados intrínsecamente motivadores. Parece que al menos algunos estados conscientes tienen de manera intrínseca la fuerza que como motivos parecen tener. En particular, los papeles funcionales y motivacionales de los estados afectivos conscientes, como los placeres y los dolores, parecen ser intrínsecos a su carácter experiencial e inseparables de sus propiedades cualitativas y fenoménicas, aunque se ha cuestionado esta opinión (Nelkin, 1989; Rosenthal, 1991). El atractivo y positivo aspecto motivacional de un placer parece ser parte de su sensación fenoménica directamente experimentada, igual que el carácter afectivo negativo de un dolor, por lo menos en el caso de la experiencia normal no patológica.

Hay un gran desacuerdo sobre cuánto se pueden disociar la sensación y la fuerza motivacional del dolor en los casos anormales, y algunos han negado la existencia de esos aspectos motivacionales intrínsecos (Dennett, 1991). Pero, por lo menos en el caso normal, la fuerza motivacional negativa del dolor *parece* estar engastada en la sensación misma de la experiencia.

Cómo ocurre esto está menos claro, y acaso sea ilusoria la apariencia de una fuerza motivacional intrínseca y experimentada directamente. Pero si es real, entonces puede ser uno de los aspectos más importantes y más antiguos evolutivamente en los que la conciencia haga una aportación a los sistemas y procesos en los que está presente (Humphreys, 1992).

Se han hecho otras propuestas sobre los posibles papeles y valores de la conciencia, y estos no agotan las posibilidades, Pero están entre las hipótesis recientes más destacadas y ofrecen un panorama razonable de los tipos de respuestas que han ofrecido a la cuestión del porqué quienes creen que la conciencia aporta algo.

6.8 Papeles constitutivos y contingentes

Hay que aclarar otro punto acerca de los varios aspectos en que las funciones propuestas pueden responder a la cuestión del porqué. En particular, hay que distinguir entre los casos *constitutivos* y los casos de realización *contingente*. En los primeros, satisfacer el papel es constitutivo de ser consciente en el sentido relevante, mientras que en los segundos la conciencia de un cierto tipo no es más que una manera, entre otras posibles, de realizar el papel requerido (Van Gulick, 1993).

Por ejemplo, poner globalmente la información a disposición de una amplia variedad de subsistemas y de aplicaciones conductuales puede ser constitutivo de la conciencia en el sentido de conciencia de acceso. En

cambio, aunque las formas cualitativa y fenoménica de la conciencia entrañen una representación muy unificada y densamente integrada de la realidad objetiva, parece posible producir representaciones que tengan estas características funcionales pero no sean de naturaleza cualitativa ni fenoménica.

El hecho de que *en nosotros* los modos de representación que tienen esas características tengan también propiedades cualitativas y fenoménicas puede que solo refleje hechos históricos contingentes relativos al diseño particular que apareció en nuestra genealogía evolutiva. Si es así, bien puede ser que haya otros medios de alcanzar un resultado comparable sin conciencia cualitativa o fenoménica. No es claro si es éste o no el modo de concebir la conciencia fenoménica y cualitativa; acaso su conexión con las representaciones unificadas y densamente integradas sea tan íntimo y constitutivo como parece serlo en el caso de la conciencia de acceso (Carruthers, 2000). Con independencia de cómo se resuelva esto, es importante no mezclar las doctrinas sobre la constitución con las doctrinas sobre la realización contingente cuando nos ocupamos de la función de la conciencia y de por qué existe (Chalmers, 1996).

7. Teorías de la conciencia

En los últimos años se han propuesto muchas teorías de la conciencia en respuesta a las cuestiones del qué, el cómo y el porqué. Pero no todas las teorías de la conciencia son teorías de la misma cosa. No solo se diferencia en los tipos de conciencia de los que se ocupan, sino en sus objetivos teóricos.

La principal división es quizá la que hay entre teorías metafísicas generales que intentan situar la conciencia en el esquema ontológico general de la realidad y teorías más específicas que ofrecen tratamientos detallados de su naturaleza, rasgos y papel. La divisoria entre ambos tipos de teorías es un poco borrosa, especialmente porque muchas teorías específicas comportan al menos algunos compromisos implícitos sobre temas metafísicos generales. Con todo, es útil recordar esa división cuando examinemos el ámbito de las propuestas teóricas.

8. Teorías metafísicas de la conciencia

Las teorías metafísicas ofrecen respuestas a la versión referida a la conciencia del problema mente-cuerpo: “¿Cuál es el estatuto ontológico de la conciencia en relación con el mundo de la realidad física?” Las respuestas ofrecidas son en gran parte paralelas a las doctrinas alternativas corrientes sobre el problema mente-cuerpo, incluyendo las principales versiones del dualismo y el fisicalismo.

8.1 Teorías dualistas

Las *teorías dualistas* consideran que al menos algunos aspectos de la conciencia caen fuera del reino de lo físico, pero las distintas formas de dualismo difieren en cuáles son esos aspectos.

El *dualismo de las sustancias*, como es el dualismo tradicional de Descartes (1644), afirma la existencia de sustancias físicas y no físicas. Estas teorías implican la existencia de mentes no físicas o yoes que son entidades a las que pertenece la conciencia. Aunque hoy día el dualismo de las sustancias goza de poco favor, tiene algunos defensores contemporáneos (Swinburne, 1986; Foster, 1989, 1996).

El *dualismo de las propiedades*, en sus diversas versiones, goza de mayor apoyo en nuestros días. Todas estas teorías sostienen la existencia de propiedades que no son idénticas ni reducibles a propiedades físicas, pero que pueden, sin embargo, ser instanciadas por las mismas cosas que instancian las propiedades físicas. En este sentido se pueden clasificar como teorías del *doble aspecto*. Admiten que ciertas partes de la realidad –organismos, cerebros, estados o procesos neuronales, instancian propiedades de dos tipos distintos y disjuntos: físicas y conscientes, fenoménicas, o cualitativas. Las teorías del doble aspecto o las dualistas pueden ser, por lo menos, de tres tipos diferentes.

El *dualismo de las propiedades fundamentales* ve las propiedades mentales conscientes como constituyentes básicos de la realidad, al mismo nivel que las propiedades físicas fundamentales, como por ejemplo, la carga electromagnética. Pueden interactuar de modos causales y legaliformes con otras propiedades fundamentales, como son las de la física, pero ontológicamente su existencia no depende ni se deriva de otras propiedades (Chalmers, 1996).

El *dualismo de las propiedades emergentes* entiende que las propiedades conscientes surgen a partir de organizaciones complejas de constituyentes físicos, pero de una manera radical, de forma que el resultado emergente está por encima de sus causas físicas y no es predecible *a priori* ni explicable en términos de sus caracteres estrictamente físicos. La coherencia de estas opiniones emergentistas ha sido cuestionada (Kim, 1998), pero tiene defensores (Hasker, 1999).

El *dualismo monista neutral de las propiedades* afirma que las propiedades mentales conscientes y las propiedades físicas dependen y derivan de un nivel más básico de realidad que no es en sí mismo mental ni físico (Russell, 1927; Strawson, 1994). Ahora bien, si el dualismo es la afirmación de que hay dos reinos distintos de entidades o propiedades fundamentales, entonces quizá no debemos clasificar el monismo neutral entre las versiones del dualismo de las propiedades, puesto que no considera que las propiedades mentales ni las físicas sean últimas o fundamentales.

El *pansiquismo* se podría interpretar como un cuarto tipo de dualismo de las propiedades porque entiende que todos los constituyentes de la realidad tienen propiedades psíquicas o por lo menos proto-psíquicas, diferentes de cualesquiera propiedades físicas que puedan tener (Nagel, 1979). Sin duda se podría combinar el monismo neutral con alguna versión del pan-proto-psiquismo (Chalmers, 1996) según el cual los aspectos proto-mentales de los micro-constituyentes pueden originar, en condiciones adecuadas, una conciencia completamente desarrollada.

No está clara la naturaleza de los aspectos proto-psíquicos relevantes, y estas teorías afrontan un dilema si es que se las propone como respuestas al Problema Difícil. O las propiedades proto-psíquicas envuelven el tipo de sensación fenoménica cualitativa que da lugar al Problema Difícil o no. Si lo hacen es difícil entender cómo podrían tener una presencia ubicua en la realidad. ¿Cómo es que los electrones o los quarks tienen sensaciones experienciales? Pero si las propiedades proto-psíquicas no entrañan esa sensación, entonces no está claro que sean más capaces que las propiedades físicas de dar cuenta de la conciencia cualitativa, ni que sirvan, pues, para resolver el Problema Difícil.

Se han dado muchos argumentos a favor del dualismo y de otras teorías anti-fisicalistas de la conciencia. Algunos son sobre todo *a priori*, como los que apelan a la supuesta inteligibilidad de los zombies (Kirk, 1970; Chalmers, 1996) o las versiones del argumento del conocimiento (Jackson, 1982, 1986), que intentan alcanzar una conclusión anti-fisicalista sobre la ontología de la conciencia a partir de los límites aparentes de nuestra capacidad para entender completamente los aspectos cualitativos de la experiencia consciente por medio de las descripciones físicas de tercera persona de los procesos cerebrales. (Véase Jackson 1998, 2004 para una opinión contraria). Hay otros argumentos a favor del dualismo que tienen motivos más empíricos, como los que apelan a las discontinuidades en las cadenas causales físicas del cerebro (Eccles y Popper, 1977) o los que se basan en supuestas anomalías en el orden temporal de la conciencia (Liber, 1982, 1985). Los argumentos dualistas de ambos tipos han sido muy cuestionados por los fisicalistas (P.S. Churchland, 1981; Dennett y Kinsbourne, 1992).

8.2 Teorías fisicalistas

La mayoría del resto de las teorías metafísicas de la conciencia son versiones de una u otra clase del fisicalismo.

Las teorías *eliminativistas* niegan de manera reduccionista la existencia de la conciencia o, por lo menos, la existencia de algunos de los tipos o rasgos de ella más comunes. Los eliminativistas radicales rechazan la noción misma de conciencia como algo confuso o erróneo y aseguran que la distinción entre lo consciente y lo no consciente no trunca por las articulaciones la realidad mental (Wilkes 1984, 1988). Piensan que la idea de conciencia se aleja tanto del blando que merece la eliminación y ser sustituida por otros conceptos y distinciones que reflejen mejor la auténtica naturaleza de la mente (P.S. Churchland, 1983).

La mayoría de los eliminativistas matizan más su negativa valoración. En vez de rechazar por completo la noción, solo objetan contra algunos rasgos prominentes que comúnmente se piensa que involucra, como los qualia (Dennett, 1990; Carruthers, 2000), el yo consciente (Dennett, 1992), o el llamado “Teatro Cartesiano” donde se proyecta internamente la sucesión temporal de la experiencia consciente (Dennett y Kinsbourne, 1992). Otros eliminativistas más modestos, como Dennett, combinan, pues, normalmente su rechazo limitado

con una teoría de los aspectos de la conciencia cuya realidad aceptan, como hace el Modelo de Los borradores Múltiples (Véase sección 9.3, más abajo).

La *teoría de la identidad*, por lo menos la teoría de la identidad psico-física de tipos, ofrece otra opción fuertemente reductiva porque identifica las propiedades, estados y procesos mentales conscientes con físicos, por lo general de carácter neuronal o neurofisiológico. Si tener una experiencia consciente cualitativa del rojo fenoménico *es meramente* estar en un estado cerebral que tenga las propiedades neurofisiológicas pertinentes, entonces esas propiedades de la experiencia son reales, pero su realidad es una realidad directamente física.

La *teoría de la identidad entre tipo y tipo* se llama así porque identifica los tipos o propiedades mentales y físicos igual que se identifica la propiedad de ser agua con la propiedad de estar compuesto por moléculas de H₂O. Tras un breve período de popularidad en los primeros tiempos del fisicalismo contemporáneo durante los años cincuenta y sesenta (Place, 1956; Smart, 1959) ha tenido poca aceptación, debido a problemas como la objeción de la realización múltiple, según la cual las propiedades mentales son más abstractas y, por lo tanto, susceptibles de ser realizadas en muchos sustratos estructurales o químicos (Fodor, 1974; Hellman y Thompson, 1975). Si la misma propiedad consciente puede ser realizada por medio de diferentes propiedades neurofisiológicas (e incluso no neurofisiológicas) en diferentes organismos, entonces las dos propiedades no pueden ser idénticas.

Con todo, la teoría de la identidad entre tipo y tipo ha gozado de un renacer reciente aunque modesto, por lo menos en lo que se refiere a los qualia o propiedades cualitativas conscientes. Eso se ha debido en parte a que hay quienes piensan que tratar la conexión psico-física relevante como una identidad ofrece una manera de disolver el problema del abismo explicativo (Hill y Mclaughlin, 1998; Papineau, 1995, 2003). Argumentan que la propiedad cualitativa consciente y la propiedad neuronal son idénticas, entonces no hay necesidad de explicar cómo la segunda explica u origina la primera: no la *causa*, sino que *es* ella. Así no hay abismo que cruzar ni hacen falta más explicaciones. Las identidades no son el tipo de cosas que se explican, pues nada es idéntico a nada, sino a sí mismo, y no tiene sentido preguntar por qué algo es idéntico a sí mismo.

Otros, sin embargo, replican que no está tan claro apelar a la identidad entre tipo y tipo excuse la necesidad de explicación (Levine, 2001). Aunque dos descripciones o concepto se refieran de hecho a una y la misma propiedad, puede ser razonable esperar una explicación de esa convergencia, alguna idea de cómo es que seleccionan una y la misma cosa aunque inicialmente o intuitivamente no parezcan tener nada que ver. En otros casos de identidades de propiedades descubiertas empíricamente, como la del calor y la energía cinética, hay una explicación que da cuenta de la convergencia de las referencias y parece apropiado esperar lo mismo en el caso de la identidad psico-física. Por lo tanto, apelar a las identidades entre tipo y tipo puede que no baste para disolver el problema del abismo explicativo.

La mayoría de las teorías de la conciencia no son eliminativistas ni se basan en identidades entre tipo y tipo. Reconocen la realidad de la conciencia pero intentan situarla en el mundo físico sobre la base de alguna relación psico-física diferente de la identidad estricta entre propiedades.

Entre las variedades comunes están las que entienden que la realidad consciente *sobreviene* a la física, está *compuesta* por la física o está *realizada* por la física.

Las teorías *funcionalistas*, en particular, se apoyan decididamente en la noción de *realización* para elucidar la relación entre la conciencia y lo físico. Según el funcionalismo, un estado o proceso es de un tipo mental o consciente en virtud del papel funcional que desempeñe en un sistema adecuadamente organizado (Block, 1980a). Un estado físico dado realiza el tipo relevante de estado mental consciente si desempeña el papel apropiado en el sistema físico que lo contiene. El funcionalista apela a menudo a analogías con otras relaciones entre niveles, como las que se dan entre lo biológico y lo bioquímico o entre lo químico y lo atómico. En cada uno de estos casos, las propiedades o hechos de un nivel son realizados por interacciones complejas entre los elementos de un nivel subyacente.

A menudo los críticos del funcionalismo niegan que la conciencia se pueda explicar adecuadamente en términos funcionales (Block 1980^a, 1980b; Levine 1983; Chalmers, 1996). Según estos críticos, puede ser que la conciencia tenga propiedades funcionales interesantes, pero su naturaleza no es esencialmente

funcional. A veces se defienden estas posturas apelando a la supuesta posibilidad de los qualia ausentes o invertidos; es decir, a la posibilidad de que haya seres equivalentes funcionalmente a los humanos normales pero que tengan los qualia invertidos o carezcan de ellos. Estas posibilidades son objeto de controversia (Shoemaker, 1981; Dennett, 1990; Carruthers, 2000), pero si se se aceptaran parece que serían un problema para el funcionalismo.

Quienes fundan un fisicalismo ontológico sobre la relación de realización, a menudo lo combinan con una visión no reduccionista al nivel conceptual o representacional que destaca la autonomía de las ciencias especiales y los distintos modos de descripción y de acceso cognitivo que proporcionan.

El *fisicalismo no reductivo* de este género niega que los recursos teóricos y conceptuales apropiados y adecuados para tratar los hechos relativos al nivel del sustrato subyacente o de la realización sean también adecuados para tratar los hechos del nivel de lo realizado (Putnam, 1975; Boyd, 1980). Como indicamos más arriba en la respuesta a la cuestión del cómo, podemos que creer que todos los hechos económicos están realizados físicamente sin por ello pensar que los recursos de las ciencias físicas ofrezcan las herramientas cognitivas y conceptuales que se necesitan para hacer economía (Fodor, 1974).

El fisicalismo no reductivo se ha sido objetado por su supuesta incapacidad para “pagar sus deudas fisicalistas” en moneda reduccionista. Su falta consiste en no dar cuenta de manera apropiada de cómo las propiedades conscientes son realizadas, o podrían serlo, por estructuras o procesos subyacentes neuronales, físicos o funcionales (Kim, 1987, 1998). Es más, se le ha acusado de incoherencia porque pretende combinar la tesis de que hay una realización física con el rechazo a explicitar esa relación de una manera inteligible estricta y *a priori* (Jackson, 2004).

Sin embargo, como se ha señalado más arriba al tratar la cuestión del cómo, los fisicalistas no reduccionistas responden aceptando que, por supuesto, es necesario dar cuenta de alguna forma de la realización psico-física, pero añaden que la explicación relevante puede no ser la deducibilidad *a priori* y, a pesar de ello, bastar para satisfacer nuestras demandas explicativas legítimas (McGinn, 1991, Van Gulick, 1985). El tema sigue sujeto a debate.

9. Teorías Específicas de la Conciencia

Si hay muchas teorías generales metafísicas u ontológicas de la conciencia, todavía más larga y diversa es la lista de las teorías detalladas sobre su naturaleza. Ni de lejos puede ser completo un breve examen de ellas, pero hay seis tipos principales de teorías que nos pueden ayudar a hacernos cargo de las posibilidades básicas: teorías de orden superior, teorías representacionales, teorías cognitivas, teorías neuronales, teorías cuánticas y teorías no físicas.

9.1 Teorías de orden superior

Las teorías de orden superior (OS) analizan la noción de estado mental consciente en términos de auto-percepción meta-mental reflexiva. El meollo de la idea es que lo que hace de un estado mental *M* un estado mental consciente es el hecho de que venga acompañado por un estado simultáneo y no inferencial de orden superior (esto es, meta-mental) cuyo contenido sea que ahora uno está en el estado *M*. Tener un deseo consciente de chocolate implica estar en dos estados mentales; hay que tener a la vez un deseo de chocolate y además estar en un estado de orden superior cuyo contenido es tener ese deseo. Los estados mentales inconscientes son inconscientes porque no tenemos, acerca de ellos, los estados relevantes de orden superior. Su condición de inconscientes consiste en el hecho de que no nos percatamos reflexiva y directamente de que estamos en ellos.

Hay dos variantes principales de las teorías de orden superior por lo que se refiere al modo psicológico de los estados meta-mentales que generan conciencia. Las teorías del pensamiento de orden superior (PSS) entienden que el estado requerido de orden superior es un meta-estado asertórico parecido a un pensamiento (Rosenthal, 1986, 1993). Las teorías de la percepción de orden superior (PRS) entienden que se parecen más a percepciones, y los asocian a algún tipo de sentido interno y de sistemas de supervisión intra-mentales (Armstrong, 1981; Lycan, 1987, 1996).

Cada una de las variantes tiene sus puntos fuertes y también sus problemas. Los teóricos de las PSS advierten que no tenemos órganos del sentido interno y aseguran que no experimentales más cualidades sensoriales que las que nos presenta la percepción directa externa. Los teóricos de las PRS, por el contrario, pueden argumentar que su postura explica algunas de las condiciones adicionales que exigen las explicaciones de OS como consecuencias naturales de la semejanza con las percepciones que tienen los estados relevantes de orden superior. En particular, la exigencia de que los meta-estados que hacen la conciencia sean no inferenciales y simultáneos con su objeto mental de nivel inferior, se podría explicar por la simultaneidad o paralelismo con que normalmente acaecen las percepciones. Percibimos lo que está ocurriendo ahora, y lo hacemos sin emplear inferencias, por lo menos no empleamos inferencias al nivel personal. Esas condiciones no son menos necesarias en las PSS, pero éstas no las explican, lo cual podría parecer que da una ventaja explicativa a las PRS (Lycan, 2004; Van Gulick, 2000), aunque algunos teóricos de las PSS argumentan otra cosa (Carruthers, 2000).

Cualesquiera que sean sus méritos respectivos, las PRS y las PSS afrontan algunas dificultades comunes, incluido lo que se puede llamar el *problema de la generalidad*. Tener un pensamiento o una percepción de un elemento dado X –sea una roca, un lápiz o una patata- no hace por lo general que X sea un X consciente. Ver o pensar en la patata que hay sobre el mostrador no hace en general que seamos conscientes de la patata. Entonces, ¿por qué tener un pensamiento o una percepción de un deseo o recuerdo dados los habría de hacer conscientes (Dretske, 1995; Byrne, 1997). Tampoco vale indicar que no aplicamos el término “consciente” a las rocas o lápices que percibimos o en los que pensamos, sino solo a los estados mentales que percibimos o en los que pensamos (Lycan, 1997; Rosenthal, 1997). Puede que eso sea cierto, pero lo que hace falta es una explicación de por qué es apropiado hacerlo.

La doctrina del orden superior es, desde luego, relevante para las formas meta-mentales de conciencia, pero algunos de sus defensores quieren que explique también otros tipos de conciencia, incluyendo los más subjetivos y los cualitativos. Una estrategia común es analizar los qualia como rasgos mentales que pueden ocurrir inconscientemente; por ejemplo se los podría explicar como propiedades de estados internos cuyas relaciones de semejanza pueden dar lugar a creencias sobre semejanzas objetivas entre elementos del mundo externo (Shoemaker, 1975, 1990). Aunque los qualia inconscientes pudieran desempeñar un papel funcional, puede que estar en un estado que tenga esos qualia no corresponda a “cómo es estar en un” estado que tenga esos qualia (Nelkin, 1989; Rosenthal, 1991, 1997). Según los teóricos del OS el *cómo-es-ser-un* solo aparece cuando nos percatamos de ese estado de primer orden y de sus propiedades cualitativas por medio del meta-estado adecuado enfocado hacia él.

Los críticos de la doctrina del OS han rechazado esa explicación, y algunos han argumentado que noción de qualia inconscientes en que se basa es incoherente (Papineau, 2002). Tengan o no éxito las explicaciones de OS de los qualia, es importante advertir que la mayoría de los defensores del OS pretenden estar ofreciendo una teoría general de la conciencia, o por lo menos el núcleo de una teoría general, no solo una teoría limitada a algunas formas especiales meta-mentales.

Otras variantes de la teoría del OS van más allá de las PSS y PRS corrientes, y algunas analizan la conciencia en términos de pensamientos disposicionales en vez de actuales (Carruthers, 2000). Otras apelan a la comprensión de orden superior implícita en vez de explícita, o eliminan o debilitan la suposición de que el meta-estado debe ser distinto y separado de su objeto de orden inferior (Gennaro, 1995; Van Gulick, 2000, 2004). Se siguen ofreciendo variantes de teorías de OS, y continúan los debates entre defensores y críticos del enfoque básico. (Véanse los últimos artículos de Gennaro, 2004).

9.2 Teorías representacionales

Casi todas las teorías de la conciencia admiten que tiene rasgos representacionales, pero las teorías representacionales se definen por la tesis más fuerte de que sus rasgos representacionales agotan sus rasgos mentales (Harman, 1990; Tye, 1995, 2000). Según los representacionistas, los estados mentales conscientes no tienen otras propiedades mentales que sus propiedades representacionales. Así, dos estados conscientes o experienciales que compartan todas sus propiedades representacionales no diferirán en ningún respecto mental.

La fuerza exacta de esta afirmación depende de cómo interpretemos la idea de ser “lo mismo representacionalmente”, para lo cual hay muchos criterios alternativos razonables. Se podría dar una definición poco refinada en términos de satisfacción o de condiciones de verdad, pero entendida de esta manera, la tesis representacionista parece claramente falsa. Hay muchas de que los estados puedan compartir las condiciones de satisfacción o de verdad aunque difieran mentalmente, incluyendo las relativas al modo de conceptualizar o presentar esas condiciones.

En el otro extremo, podríamos contar dos estados como distintos representacionalmente si difieren en algún rasgo que desempeñe un papel en su función u operación representacional. Bajo esta interpretación tan amplia, cuales quiera diferencias entre los portadores de contenido contarían como diferencias representacionales, aunque tengan el mismo contenido intencional o representacional; podrían diferir en los *medios* o en el *modo* de representación, no en el *contenido*.

Bajo esta interpretación crecería, desde luego, la plausibilidad de la idea de que las propiedades representacionales de un estado consciente agotan sus propiedades mentales, pero a costa de debilitar mucho o incluso trivializar la tesis. Por lo tanto, el representacionista parece necesitar una interpretación de *identidad representacional* que vaya más allá de las meras condiciones de satisfacción y refleje todos los aspectos intencionales o de contenido de la representación sin que sea sensibles las meras diferencias en los rasgos subyacentes ajenos al contenido de los procesos del nivel de la realización. Así, la mayoría de los representacionistas proponen condiciones de la experiencia consciente que incluyen una condición de contenido junto con requisitos sobre el papel causal o el formato (Tye, 1995; Dretske, 1995; Carruthers, 2000). Otros representacionistas aceptan la existencia de los qualia, pero los tratan como propiedades objetivas con que se representan los objetos externos, esto es, las tratan como *propiedades representadas*, no como *propiedades de las representaciones* o como estados mentales (Dretske, 1995; Lycan, 1996).

Se puede entender el representacionismo como una forma especial de eliminativismo, puesto que niega la existencia de las propiedades que normalmente pensamos que tienen los estados mentales conscientes; a saber, las que son mentales, pero no representacionales. Parece que los qualia, por lo menos si los entendemos como propiedades monádicas intrínsecas accesibles a la introspección de los estados conscientes, parecen los objetivos más claros de esa eliminación. En efecto, parte del motivo para sostener el representacionismo es mostrar que podemos acomodar todos los hechos relativos a la conciencia, quizá en un marco fisicalista, sin tener que hacer sitio para los qualia o cualesquiera otras propiedades mentales que a primera vista no sean representacionales (Dennett, 1990; Lycan, 1996; Carruthers, 2000).

El representacionismo ha sido muy popular en los últimos tiempos y ha tenido muchos defensores, pero sigue estando muy controvertido y tropieza con nuestras intuiciones sobre casos y experimentos mentales clave (Block, 1996). En particular, la posibilidad de los qualia invertidos ofrece un experimento crucial. Para los contrarios al representacionismo la mera posibilidad lógica de los qualia invertidos muestra que los estados conscientes pueden diferir en un aspecto mental significativo y, sin embargo, coincidir desde el punto de vista representacional. Los representacionistas, por el contrario, niegan la posibilidad de esa inversión o su supuesta importancia (Dretske, 1995; Tye, 2000).

A favor y en contra del representacionismo se han propuesto otros argumentos, como los relativos a las percepciones en diferentes modalidades sensoriales de un mismo estado de cosas -ver y tocar el mismo cubo que, según parece, podrían implicar diferencias mentales distintas de cómo los estados relevantes representan que es el mundo (Peacocke, 1983; Tye, 2003). En cada caso ambas partes pueden aportar fuertes intuiciones e ingeniosos argumentos. El debate continúa.

9.3 Teorías cognitivas

Los filósofos y los psicólogos han ofrecido modelos que pretenden explicar la conciencia en términos de procesos cognitivos.

El ejemplo filosófico más destacado es el Modelo de los Borradores Múltiples (MBM) de la conciencia, propuesto por Daniel Dennett (1991). Combina elementos del representacionalismo y de la teoría del orden superior, pero lo hace de modo que se diferencia curiosamente de las versiones más corrientes de ambos. El MBM combina muchos rasgos distintos pero relacionados.

Su nombre refleja el hecho de que en cualquier momento dado ocurren en el cerebro muchas determinaciones del contenido. Lo que hace que algunas de éstas sean conscientes no es que ocurran en un lugar privilegiado espacial o funcional –el llamado “Teatro Cartesiano”– ni un modo o formato especiales. Todo esto lo niega la MBM. Es más bien cuestión de lo que Dennett llama “celebridad cerebral”, esto es, del grado en que un contenido dado influye el desarrollo futuro de otros contenidos en el cerebro, especialmente con vistas a cómo esos efectos se manifiestan en los informes y conductas que la persona hace en respuesta a distintos “sondeos” para indicar su estado consciente.

Una de las afirmaciones clave del MBM es que diferentes sondeos (por ejemplo, que a uno le hagan diferentes preguntas o estar en diferentes contextos que planteen demandas conductuales diferentes) pueden obtener diferentes respuestas sobre el estado consciente de la persona. Más aun, según el MBM puede que no haya hechos independientes del sondeo acerca de cuál es el estado consciente en el que la persona está realmente. De ahí el “múltiples” del Modelo de los Borradores Múltiples.

El MBM es representacionista porque analiza la conciencia en términos de las relaciones de contenido. También niega la existencia de los qualia, y así rechaza cualquier intento de distinguir por su presencia entre los estados conscientes y los no conscientes. También rechaza la noción del yo como observador interno, en el Teatro Cartesiano o donde sea. El MBM trata el yo como un aspecto emergente o virtual de la narración coherente, más o menos serial, que se construye por medio del juego interactivo de los contenidos del sistema. Muchos de esos contenidos están atados juntos en el nivel intencional como percepciones o determinaciones desde un punto de vista relativamente unificado y temporalmente extendido; esto es, sus contenidos son coherentes, como si fueran las experiencias de un yo continuo. Pero lo crucial para el MBM es el orden de la dependencia. Los contenidos relevantes no están unificados porque los observe un yo singular, sino al revés. Como están unificados y son coherentes en el nivel del contenido, cuentan como experiencias de un yo singular, o por lo menos de un yo singular virtual.

Es en este aspecto en el que el MBM comparte algunos elementos con las teorías del orden superior. Los contenidos que componen la narración serial son, por lo menos implícitamente, los de un yo continuo, incluso aunque sea virtual, y son ellos los que tienen más posibilidades de ser expresados en los informes que la persona hace sobre su estado consciente en diversos sondeos. Ellos implican, pues, cierto grado de reflexividad o auto-percatación del tipo que es básico para las teorías del orden superior, pero ese aspecto del orden superior es más bien un rasgo implícito del flujo de contenidos que algo presente en estados explícitos de orden superior del tipo que se encuentran en las teorías OS corrientes.

Un ejemplo psicológico destacado del enfoque cognitivo es la teoría del espacio global de trabajo de Bernard Baars (1988). Ofrece un modelo psicológico y en gran medida funcional de la conciencia, que trata directamente la noción de conciencia de acceso y tiene mucho en común con la idea de la celebridad cerebral del modelo de los borradores múltiples. La idea principal de las teorías del espacio global de trabajo es que la conciencia es una capacidad o módulo con recursos limitados que permite que la información se “emita” a través del sistema y permite un procesamiento más flexible y refinado. De este modo, tiene un parentesco cercano con muchos modelos de la psicología cognitiva que se refieren a la atención y a la memoria de trabajo. Se han propuesto otros muchos modelos cognitivos de la conciencia de carácter psicológico, incluyendo los de George Mandler (1975), Timothy Shallice (1988) y Daniel Schacter (1989).

9.4 Teorías neuronales

Hay muchas formas de teorías neuronales, aunque la mayoría se ocupan de uno u otro modo de los llamados “correlatos neuronales de la conciencia” o CNC. Salvo que seamos dualistas o, en general, no-fisicalistas, se requiere algo más que la mera correlación; algunos de los CNC tienen que ser los sustratos esenciales de la conciencia. Una teoría neuronal explicativa tiene que explicar por qué o cómo se dan las correlaciones relevantes, y si la teoría está comprometida con el fisicalismo, eso exigirá mostrar cómo los sustratos neuronales pueden ser los sustratos neuronales de la conciencia o, por lo menos, cómo pueden realizarlos satisfaciendo los papeles o condiciones requeridos (Metzinger, 2000).

Esas teorías no solo se diferencian por los procesos o propiedades neuronales a los que apelan, sino por los aspectos de la conciencia que toman como *explananda*. Unas se basan en rasgos sistémicos de alto nivel del cerebro, y otras se centran en propiedades fisiológicas o estructurales más específicas, con las diferencias correspondientes en sus objetivos explicativos.

Un muestrario de las teorías neuronales recientes incluye los modelos que apelan a campos globales integrados (Kinsbourne), la ligadura mediante la oscilación sincrónica (Singer, 1999; Crick y Koch, 1990), las asambleas neuronales mediadas por el NMDA (Floh, 1995), los patrones de activación cortical modulados por el tálamo (Llinás, 2001), bucles corticales reentrantes (Edelman, 1989), mecanismos comparadores que se involucran en bucles continuos de acción-predicción, valoración entre las áreas frontal y media del cerebro (Gray, 1995), procesos interpretativos situados en el hemisferio izquierdo (Gazzaniga, 1988) y procesos emotivos somato-sensoriales hemostáticos basados en el nexo frontal-límbico (Damasio, 1999) o en la sustancia gris periacueductal (Panksepp, 1998).

En cada caso el propósito es explicar cómo la organización y la actividad del nivel neuronal adecuado pueden estar por debajo de uno u otro tipo o rasgo de la conciencia. Los campos globales o las asambleas sincrónicas transitorias podrían subyacer a la unidad intencional de la conciencia fenoménica. La plasticidad basada en el NMDA, las proyecciones talámicas específicas sobre la corteza o las ondas regulares oscilatorias podrían contribuir a la formación de regularidades o patrones de breve duración, pero extendidos, necesarios para unir la experiencia consciente integrada a partir de la actividad local de diversos módulos cerebrales especializados. Los procesos interpretativos del hemisferio izquierdo podrían suministrar la base para las formas narrativas de la auto-percepción consciente. Así, es posible que muchas teorías neuronales distintas sean todas ellas verdaderas, y que cada una contribuya a una comprensión parcial de las conexiones entre la mente consciente en sus formas diversas y el cerebro activo en sus muchos niveles de organización y de estructura.

9.5 Teorías cuánticas

Hay otras teorías físicas que han ido más allá de lo neuronal y han colocado el lugar natural de la conciencia a un nivel mucho más fundamental, en el nivel microfísico de los fenómenos cuánticos. Según esas teorías, la naturaleza y la base de la conciencia no se pueden explicar en el marco de la física clásica, sino que se han de buscar en la imagen alternativa de la realidad física que da la mecánica cuántica. Quienes proponen el enfoque cuántico para la conciencia entienden que el carácter radicalmente alternativo, y muchas veces contrario a la intuición, de la física cuántica es justo lo que hace falta para superar los supuestos obstáculos explicativos que afrontan los intentos más ordinarios de pasar el abismo psico-físico.

También aquí se ha propuesto un amplio abanico de teorías y modelos específicos que apelan a una variedad de fenómenos cuánticos para explicar distintos rasgos de la conciencia. Aquí sería imposible catalogarlos o explicar siquiera un poco de la sustancia de los rasgos clave de la mecánica cuántica a los que apelan. Sin embargo, un repaso breve y selectivo acaso dé una idea, aunque sea parcial y oscura, de las posibilidades que se han propuesto. El físico Roger Penrose (1989, 1994) y el anestesiólogo Stuart Hameroff (1998) han abanderado un modelo según el cual la conciencia surge gracias a efectos cuánticos que ocurren en las estructuras subcelulares internas de las neuronas conocidas como *microtúbulos*. El modelo postula los llamados “colapsos objetivos” que involucra el sistema cuántico cuando pasa de una superposición de múltiples estados posibles a un único estado definido, pero sin la intervención de un observador o una medición, como la mayoría de los modelos mecánicos cuánticos. Según Penrose y Hameroff el entorno interno de los microtúbulos es especialmente apropiado para tales colapsos objetivos, y los auto-colapsos

resultantes producen un flujo coherente de actividad neuronal reguladora y posibilitan procesos mentales no algorítmicos.

El psiquiatra Ian Marshall ha ofrecido un modo que intenta explicar unidad coherente de la conciencia apelando a la producción en el interior del cerebro de un estado físico semejante al de un *condensado de Bose-Einstein*. Este es un fenómeno cuántico en el que una colección de átomos actúa como una única unidad coherente y se pierde la distinción entre los átomos discretos. Aunque los estados cerebrales no son literalmente ejemplos de condensados Bose-Einstein, se han ofrecido razones para mostrar por qué los cerebros son capaces de originar estados que pueden mostrar una coherencia similar (Marshall&Zohar, 1990)

También se ha buscado una base para la conciencia en la naturaleza *holística* de la mecánica cuántica y en el fenómeno del *entrelazamiento*, según el cual las partículas que han interactuado continúan dependiendo unas de otras después de haberse separado. No es de extrañar que se hayan usado estos modelos sobre todo para explicar la coherencia de la conciencia, pero también se han aducido como objeción general a la concepción atomística de la física general según la cual las propiedades de las totalidades se deben explicar por medio de las propiedades de sus partes junto con el modo de combinación, método de explicación que hasta la fecha, se puede pensar, se ha mostrado incapaz de explicar la conciencia (Silberstein 1998, 2001).

Otros han pensado que la mecánica cuántica indica que la conciencia es una propiedad absolutamente fundamental de la realidad física, y que hay que buscarla en el nivel más básico (Stapp, 1993) Han apelado especialmente al papel del observador en el colapso de la función de onda, esto es, el colapso de la realidad cuántica, a partir de la superposición de estados posibles, en un estado definido único cuando se hace una medición. Esos modelos pueden adoptar o no una forma de cuasi-idealismo, en el que la misma existencia de la realidad física depende de que sea observada conscientemente.

Se pueden encontrar otros muchos modelos cuánticos de la conciencia, algunos de los cuales abogan por una metafísica radicalmente revisionistas, aunque otros no, pero estos cuatro ofrecen una muestra razonable, aunque parcial, de las alternativas.

9.6 Teorías no físicas

La mayoría de las teorías específicas de la conciencia –cognitivas, neuronales o cuánticas- intentan explicar o dar modelos de la conciencia como rasgo natural del mundo físico. Pero quienes rechazan una ontología física de la conciencia tienen que encontrar maneras de teorizarla como aspecto no físico de la realidad. Así que quienes adoptan una metafísica dualista o anti-fisicalista, al final tienen que proporcionar modelos específicos de la conciencia que difieran de los cinco tipos anteriores. Los dualistas de las sustancias y de las propiedades tienen que desarrollar los detalles de sus teorías para articular los rasgos no físicos relevantes de la realidad con los que identifican la conciencia o a los quieren apelar para explicarla.

Se han propuesto diversos modelos, entre ellos los siguientes. David Chalmers (1996) ha ofrecido una versión, que concede que es especulativa, del pampsiquismo que apela a la noción de información no solo para explicar las invariaciones psico-físicas entre los espacios de la información fenoménica y de la realizada físicamente, sino también para explicar la ontología de lo físico a partir de lo informacional (una versión de la teoría “*it from bit*”). En una línea algo parecida, Gregg Rosenberg ha propuesto (2004) una explicación de la conciencia que aborda simultáneamente la última base categorial de las relaciones causales. En el caso de la causalidad, como en el de la conciencia, Rosenberg argumenta que los hechos relacionales y funcionales tienen que depender finalmente de una base categórica no relacional, y ofrece un modelo según el cual las relaciones causales y los hechos cualitativos fenoménicos dependen de la misma base. Además, como hemos indicado más arriba (sec. 9.5), algunas teorías cuánticas tratan la conciencia como un rasgo fundamental de la realidad (Stapp, 1993), y al hacer esto también se pueden clasificar razonablemente como teorías no físicas

10. Conclusión

Lo más verosímil es un completo entendimiento de la conciencia requiera teorías de muchos tipos. Podría ser útil, sin entrañar contradicción, aceptar diversos modelos que traten de explicar respectivamente los aspectos de la conciencia físico, neuronal, cognitivo, funcional, representacional y de orden superior. No es verosímil que una sola perspectiva teórica baste para explicar todos los rasgos de la conciencia que queremos entender. Por lo tanto, es posible que el mejor camino para el progreso futuro lo ofrezca una aproximación sintética y pluralista.

Bibliografia

- Akins, K. 1993. "A bat without qualities?" In M. Davies and G. Humphreys, eds. *Consciousness: Psychological and Philosophical Essays*. Oxford: Blackwell.
- Akins, K. 1996. "Lost the plot? Reconstructing Dennett's multiple drafts theory of consciousness." *Mind and Language*, 11: 1-43.
- Anderson, J. 1983. *The Architecture of Cognition*. Cambridge, MA: Harvard University Press.
- Armstrong, D. 1968. *A Materialist Theory of Mind*, London: Routledge and Kegan Paul.
- Armstrong, D. 1981. "What is consciousness?" In *The Nature of Mind*. Ithaca, NY: Cornell University Press.
- Baars, B. 1988. *A Cognitive Theory of Consciousness*. Cambridge: Cambridge University Press.
- Balog, K. 1999. "Conceivability, possibility, and the mind-body problem." *Philosophical Review*, 108: 497-528.
- Block, N. 1980a. "Troubles with Functionalism," in *Readings in the Philosophy of Psychology*, Volume 1, Ned Block, ed., Cambridge, MA : Harvard University Press, 268-305.
- Block, N. 1980b. "Are absent qualia impossible?" *Philosophical Review*, 89/2: 257-74.
- Block, N. 1990. "Inverted Earth," *Philosophical Perspectives*, 4, J. Tomberlin, ed., Atascadero, CA: Ridgeview Publishing Company.
- Block, N. 1995. "On a confusion about the function of consciousness." *Behavioral and Brain Sciences*, 18: 227-47.
- Block, N. 1994. "What is Dennett's theory a theory of?" *Philosophical Topics*, 22/1-2: 23-40.
- Block, N. 1996. "Mental paint and mental latex." In E. Villanueva, ed. *Perception*. Atascadero, CA: Ridgeview.
- Block, N. and Stalnaker, R. 1999. "Conceptual analysis, dualism, and the explanatory gap." *Philosophical Review*, 108/1: 1-46.
- Boyd, R. 1980. "Materialism without reductionism: What physicalism does not entail." In N. Block, ed. *Readings in the Philosophy of Psychology*, Vol. 1. Cambridge, MA: Harvard University Press.
- Byrne, A. 1997. "Some like it HOT: consciousness and higher-order thoughts." *Philosophical Studies*, 2: 103-29.
- Byrne, A. 2001. "Intentionalism defended". *Philosophical Review*, 110: 199-240.
- Campbell, K. 1970. *Body and Mind*. New York: Doubleday.
- Campbell, J. 1994. *Past, Space, and Self*. Cambridge, MA: MIT Press.
- Carruthers, P. 2000. *Phenomenal Consciousness*. Cambridge: Cambridge University Press.
- Chalmers, D. 1995. "Facing up to the problem of consciousness". *Journal of Consciousness Studies*, 2: 200-19.
- Chalmers, D. 1996. *The Conscious Mind*. Oxford: Oxford University Press.
- Chalmers, D. 2002. "Does conceivability entail possibility?" In T. Gendler and J. Hawthorne eds. *Conceivability and Possibility*. Oxford: Oxford University Press.
- Chalmers, D. 2003. "The content and epistemology of phenomenal belief." In A. Jokic and Q. Smith eds. *Consciousness: New Philosophical Perspectives*. Oxford: Oxford University Press.
- Chalmers, D. and Jackson, F. 2001. "Conceptual analysis and reductive explanation". *Philosophical Review*, 110/3: 315-60.
- Churchland, P. M. 1985. "Reduction, qualia, and direct introspection of brain states". *Journal of Philosophy*, 82: 8-28.
- Churchland, P. M. 1995. *The Engine of Reason and Seat of the Soul*. Cambridge, MA: MIT Press.
- Churchland, P. S. 1981. "On the alleged backwards referral of experiences and its relevance to the mind body problem". *Philosophy of Science*, 48: 165-81.
- Churchland, P. S. 1983. "Consciousness: the transmutation of a concept". *Pacific Philosophical Quarterly*, 64: 80-95.
- Churchland, P. S. 1996. "The hornswoggle problem". *Journal of Consciousness Studies*, 3: 402-8.
- Clark, A. 1993. *Sensory Qualities*. Oxford: Oxford University Press.
- Clark, G. and Riel-Salvatore, J. 2001. "Grave markers, middle and early upper paleolithic burials". *Current Anthropology*, 42/4: 481-90.
- Cleeremans, A., ed. 2003. *The Unity of Consciousness: Binding, Integration and Dissociation*. Oxford: Oxford University Press.
- Crick, F. and Koch, C. 1990. "Toward a neurobiological theory of consciousness". *Seminars in Neuroscience*, 2: 263-75.
- Crick, F. H. 1994. *The Astonishing Hypothesis: The Scientific Search for the Soul*. New York: Scribners.
- Davies, M. and Humphreys, G. 1993. *Consciousness: Psychological and Philosophical Essays*. Oxford: Blackwell.
- Damasio, A. 1999. *The Feeling of What Happens: Body and Emotion in the Making of Consciousness*. New York: Harcourt.

- Dennett, D. C. 1978. *Brainstorms*. Cambridge: MIT Press.
- Dennett, D. C. 1984. *Elbow Room: The Varieties of Free Will Worth Having*. Cambridge: MIT Press.
- Dennett, D. C. 1990. "Quining qualia". In *Mind and Cognition*, W. Lycan, ed., Oxford : Blackwell, 519-548.
- Dennett, D. C. 1991. *Consciousness Explained*. Boston: Little, Brown and Company.
- Dennett, D. C. 1992. "The self as the center of narrative gravity". In F. Kessel, P. Cole, and D. L. Johnson, eds. *Self and Consciousness: Multiple Perspectives*. Hillsdale, NJ: Lawrence Erlbaum.
- Dennett, D. C. 2003. *Freedom Evolves*. New York: Viking.
- Dennett, D. C. and Kinsbourne, M. 1992. "Time and the observer: the where and when of consciousness in the brain". *Behavioral and Brain Sciences*, 15: 187-247.
- Descartes, R. 1644/1911. *The Principles of Philosophy*. Translated by E. Haldane and G. Ross. Cambridge: Cambridge University Press.
- Dretske, F. 1993. "Conscious experience." *Mind*, 102: 263-283.
- Dretske, F. 1994. "Differences that make no difference". *Philosophical Topics*, 22/1-2: 41-58.
- Dretske, F. 1995. *Naturalizing the Mind*. Cambridge, Mass: The MIT Press, Bradford Books.
- Eccles, J. and Popper, K. 1977. *The Self and Its Brain: An Argument for Interactionism*. Berlin: Springer
- Edelman, G. 1989. *The Remembered Present: A Biological Theory of Consciousness*. New York: Basic Books.
- Farah, M. 1990. *Visual Agnosia*. Cambridge: MIT Press.
- Flanagan, O. 1992. *Consciousness Reconsidered*. Cambridge, MA: MIT Press.
- Flohr, H. 1995. "An information processing theory of anesthesia". *Neuropsychologia*, 33/9: 1169-80.
- Flohr, H., Glade, U. and Motzko, D. 1998. "The role of the NMDA synapse in general anesthesia". *Toxicology Letters*, 100-101: 23-29.
- Fodor, J. 1974. "Special sciences". *Synthese*, 28: 77-115.
- Fodor, J. 1983. *The Modularity of Mind*. Cambridge, MA: MIT Press.
- Foster, J. 1989. "A defense of dualism". In J. Smythies and J. Beloff, eds. *The Case for Dualism*. Charlottesville, VA: University of Virginia Press.
- Foster J. 1996. *The Immaterial Self: A Defence of the Cartesian Dualist Conception of Mind*. London: Routledge.
- Gallistel, C. 1990. *The Organization of Learning*. Cambridge, MA: MIT Press.
- Gardiner, H. 1985. *The Mind's New Science*. New York: Basic Books.
- Gazzaniga, M. 1988. *Mind Matters: How Mind and Brain Interact to Create our Conscious Lives*. Boston: Houghton Mifflin.
- Gennaro, R. 1995. *Consciousness and Self-consciousness: A Defense of the Higher-Order Thought Theory of Consciousness*. Amsterdam and Philadelphia: John Benjamins.
- Gennaro, R., ed. 2004. *Higher-Order Theories of Consciousness*. Amsterdam and Philadelphia: John Benjamins.
- Gray, J. 1995. "The contents of consciousness: a neuropsychological conjecture". *Behavior and Brain Sciences*, 18/4: 659-722.
- Hameroff, S. 1998. "Quantum computation in brain microtubules? The Penrose-Hameroff "Orch OR" model of consciousness". *Philosophical Transactions Royal Society London*, A 356: 1869-96.
- Hardin, C. 1986. *Color for Philosophers*. Indianapolis: Hackett.
- Hardin, C. 1992. "Physiology, phenomenology, and Spinoza's true colors". In A. Beckermann, H. Flohr, and J. Kim, eds. *Emergence or Reduction?: Prospects for Nonreductive Physicalism*. Berlin and New York: De Gruyter.
- Harman, G. 1990. "The intrinsic quality of experience". In J. Tomberlin, ed. *Philosophical Perspectives*, 4. Atascadero, CA: Ridgeview Publishing .
- Hartshorne, C. 1978. "Panpsychism: mind as sole reality". *Ultimate Reality and Meaning*, 1: 115-29.
- Hasker, W. 1999. *The Emergent Self*. Ithaca, NY: Cornell University Press.
- Heidegger, M. 1927/1962. *Being and Time (Sein und Zeit)*. Translated by J. Macquarrie and E. Robinson. New York: Harper and Row.
- Hellman, G. and Thompson, F. 1975. "Physicalism: ontology, determination and reduction". *Journal of Philosophy*, 72: 551-64.
- Hill, C. 1991. *Sensations: A Defense of Type Materialism*. Cambridge: Cambridge University Press.
- Hill, C. 1997. "Imaginability, conceivability, possibility, and the mind-body problem". *Philosophical Studies*, 87: 61-85.
- Hill, C. and McLaughlin, B. 1998. "There are fewer things in reality than are dreamt of in Chalmers' philosophy". *Philosophy and Phenomenological Research*, 59/2: 445-54.
- Horgan, T. 1984. "Jackson on Physical Information and Qualia." *Philosophical Quarterly*, 34: 147-83.

- Horgan, T. and Tienson, J. 2002. "The intentionality of phenomenology and the phenomenology of intentionality". In D. J. Chalmers, ed., *Philosophy of Mind: Classical and Contemporary Readings*. New York: Oxford University Press.
- Hume, D. 1739/1888. *A Treatise of Human Nature*. ed. L. Selby-Bigge. Oxford: Oxford University Press.
- Humphreys, N. 1982. *Consciousness Regained*. Oxford: Oxford University Press.
- Humphreys, N. 1992. *A History of the Mind*. London: Chatto and Windus.
- Husserl, E. 1913/1931. *Ideas: General Introduction to Pure Phenomenology (Ideen au einer reinen Phänomenologie und phänomenologischen Philosophie)*. Translated by W. Boyce Gibson. New York: MacMillan.
- Husserl, E. 1929/1960. *Cartesian Meditations: an Introduction to Phenomenology*. Translated by Dorian Cairns. The Hague: M. Nijhoff.
- Huxley, T. 1866. *Lessons on Elementary Physiology* 8. London
- Huxley, T. 1874. "On the hypothesis that animals are automata". *Fortnightly Review*, 95: 555-80. Reprinted in *Collected Essays*. London, 1893.
- Hurley, S. 1998. *Consciousness in Action*. Cambridge, MA: Harvard University Press.
- Jackson, F. 1982. "Epiphenomenal qualia". *Philosophical Quarterly*, 32: 127-136.
- Jackson, F. 1986. "What Mary didn't know". *Journal of Philosophy*, 83: 291-5.
- Jackson, F. 1993. "Armchair metaphysics". In J. O'Leary-Hawthorne and M. Michael, eds. *Philosophy of Mind*. Dordrecht: Kluwer Books.
- Jackson, F. 1998. "Postscript on qualia". In F. Jackson *Mind, Method and Conditionals*. London: Routledge.
- Jackson, F. 2004. "Mind and illusion." In P. Ludlow, Y. Nagasawa and D. Stoljar eds. *There's Something about Mary: Essays on the Knowledge Argument*. Cambridge, MA: MIT Press.
- James, W. 1890. *The Principles of Psychology*. New York: Henry Holt and Company.
- Jaynes, J. 1974. *The Origins of Consciousness in the Breakdown of the Bicameral Mind*. Boston: Houghton Mifflin.
- Kim, J. 1987. "The myth of non-reductive physicalism". *Proceedings and Addresses of the American Philosophical Association*.
- Kim, J. 1998. *Mind in Physical World*. Cambridge: MIT Press.
- Kant, I. 1787/1929. *Critique of Pure Reason*. Translated by N. Kemp Smith. New York: MacMillan.
- Kinsbourne, M. 1988. "Integrated field theory of consciousness". In A. Marcel and E. Bisiach, eds. *Consciousness in Contemporary Science*. Oxford: Oxford University Press.
- Kirk, R. 1974. "Zombies vs materialists". *Proceedings of the Aristotelian Society*, Supplementary Volume, 48: 135-52
- Kirk, R. 1991. "Why shouldn't we be able to solve the mind-body problem?" *Analysis*, 51: 17-23.
- Köhler, W. 1929. *Gestalt Psychology*. New York: Liveright.
- Köffka, K. 1935. *Principles of Gestalt Psychology*. New York: Harcourt Brace.
- Leibniz, G. W. 1686 /1991. *Discourse on Metaphysics*. Translated by D. Garter and R. Aries. Indianapolis: Hackett.
- Leibniz, G. W. 1720/1925. *The Monadology*. Translated by R. Lotte. London: Oxford University Press.
- Levine, J. 1983. "Materialism and qualia: the explanatory gap". *Pacific Philosophical Quarterly*, 64: 354-361.
- Levine, J. 1993. "On leaving out what it's like". In M. Davies and G. Humphreys, eds. *Consciousness: Psychological and Philosophical Essays*. Oxford: Blackwell.
- Levine, J. 1994. "Out of the closet: a qualophile confronts qualophobia". *Philosophical Topics*, 22/1-2: 107-26.
- Levine, J. 2001. *Purple Haze: The Puzzle of Conscious Experience*. Cambridge, Mass: The MIT Press.
- Lewis, D. 1972. "Psychophysical and theoretical identifications". *Australasian Journal of Philosophy*, 50: 249-58.
- Lewis, D. 1990. "What experience teaches." In W. Lycan, ed. *Mind and Cognition: A Reader*. Oxford: Blackwell.
- Libet, B. 1982. "Subjective antedating of a sensory experience and mind-brain theories". *Journal of Theoretical Biology*, 114: 563-70.
- Libet, B. 1985. "Unconscious cerebral initiative and the role of conscious will in voluntary action". *Behavioral and Brain Sciences*, 8: 529-66.
- Llinas, R. 2001. *I of the vortex: from neurons to self*. Cambridge, MA: MIT Press
- Loar, B. 1990. "Phenomenal states," in *Philosophical Perspectives*, 4: 81-108.
- Loar, B. 1997. "Phenomenal states". In N. Block, O. Flanagan, and G. Guzeldere eds. *The Nature of Consciousness*. Cambridge, MA: MIT Press.
- Locke, J. 1688/1959. *An Essay on Human Understanding*. New York: Dover.
- Lockwood, M. 1989. *Mind, Brain, and the Quantum*. Oxford: Oxford University Press.

- Lorenz, K. 1977. *Behind the Mirror (Rückseite dyes Speigels)*. Translated by R. Taylor. New York: Harcourt Brace Jovanovich.
- Lycan, W. 1987. *Consciousness*. Cambridge, MA: MIT Press.
- Lycan, W. 1996. *Consciousness and Experience*. Cambridge, MA: MIT Press.
- Lycan, W. 2004. "The superiority of HOP to HOT". In R. Gennaro ed. *Higher-Order Theories of Consciousness*. Amsterdam and Philadelphia: John Benjamins.
- Mandler, G. 1975. *Mind and Emotion*. New York: Wiley.
- Marshall, I. and Zohar, D. 1990. *The Quantum Self: Human Nature and Consciousness Defined by the New Physics*. New York: Morrow.
- McGinn, C. 1989. "Can we solve the mind-body problem?" *Mind*, 98: 349-66
- McGinn, C. 1991. *The Problem of Consciousness*. Oxford: Blackwell.
- McGinn, C. 1995. "Consciousness and space." In T. Metzinger, ed. *Conscious Experience*. Paderborn: Ferdinand Schöningh.
- Merleau-Ponty, M. 1945/1962. *Phenomenology of Perception (Phénoménologie de lye Perception)*. Translated by Colin Smith. London: Routledge and Kegan Paul.
- Metzinger, T., ed. 1995. *Conscious Experience*. Paderborn: Ferdinand Schöningh.
- Metzinger, T. ed. 2000. *Neural Correlates of Consciousness: Empirical and Conceptual Questions*. Cambridge, MA: MIT Press.
- Mill, J. 1829. *Analysis of the Phenomena of the Human Mind*. London.
- Mill, J.S. 1865. *An Analysis of Sir William Hamilton's Philosophy*. London.
- Moore, G. E. 1922. "The refutation of idealism." In G. E. Moore *Philosophical Studies*. London : Routledge and Kegan Paul.
- Nagel, T. 1974. "What is it like to be a bat?" *Philosophical Review*, 83: 435-456.
- Nagel, T. 1979. "Panpsychism." In T. Nagel *Mortal Questions*. Cambridge: Cambridge University Press.
- Natsoulas, T. 1983. "Concepts of consciousness." *Journal of Mind and Behavior*, 4: 195-232.
- Nelkin, N. 1989. "Unconscious sensations." *Philosophical Psychology*, 2: 129-41.
- Nemirow, L. 1990. "Physicalism and the cognitive role of acquaintance." In W. Lycan, ed. *Mind and Cognition: A Reader*. Oxford: Blackwell.
- Neisser, U. 1965. *Cognitive Psychology*. Englewood Cliffs: Prentice Hall.
- Nida-Rümelin, M. 1995. "What Mary couldn't know: belief about phenomenal states." In T. Metzinger, ed. *Conscious Experience*. Paderborn: Ferdinand Schöningh.
- Panksepp, J. 1998. *Affective Neuroscience*. Oxford: Oxford University Press.
- Papineau, D. 1994. *Philosophical Naturalism*. Oxford: Blackwell.
- Papineau, D. 1995. "The antipathetic fallacy and the boundaries of consciousness." In T. Metzinger, ed. *Conscious Experience*. Paderborn: Ferdinand Schöningh.
- Papineau, D. 2002. *Thinking about Consciousness*. Oxford: Oxford University Press.
- Peacocke, C. 1983. *Sense and Content*, Oxford: Oxford University Press.
- Pearson, M.P. 1999. *The Archeology of Death and Burial*. College Station, Texas: Texas A&M Press.
- Penfield, W. 1975. *The Mystery of the Mind: a Critical Study of Consciousness and the Human Brain*. Princeton, NJ: Princeton University Press.
- Perry, J. 2001. *Knowledge, Possibility, and Consciousness*. Cambridge, MA: MIT Press.
- Penrose, R. 1989. *The Emperor's New Mind: Computers, Minds and the Laws of Physics*. Oxford: Oxford University Press.
- Penrose, R. 1994. *Shadows of the Mind*. Oxford: Oxford University Press.
- Place, U. T. 1956. "Is consciousness a brain process?" *British Journal of Psychology*, 44-50.
- Putnam, H. 1975. "Philosophy and our mental life." In H. Putnam *Mind Language and Reality: Philosophical Papers Vol. 2*. Cambridge: Cambridge University Press.
- Putnam, H. and Oppenheim, P. 1958. "Unity of science as a working hypothesis." In H. Fiegl, G. Maxwell, and M. Scriven eds. *Minnesota Studies in the Philosophy of Science II*. Minneapolis: University of Minnesota Press.
- Rey, G. 1986. "A question about consciousness." In H. Otto and J. Tuedio, eds. *Perspectives on Mind*. Dordrecht: Kluwer.
- Robinson, H. 1982. *Matter and Sense: A Critique of Contemporary Materialism*. Cambridge: Cambridge University Press.
- Robinson, D. 1993. "Epiphenomenalism, laws, and properties." *Philosophical Studies*, 69: 1-34.

- Rosenberg, G. 2004. *A Place for Consciousness: Probing the Deep Structure of the Natural World*. New York: Oxford University Press.
- Rosenthal, D. 1986. "Two concepts of consciousness." *Philosophical Studies*, 49: 329-59.
- Rosenthal, D. 1991. "The independence of consciousness and sensory quality." In E. Villanueva, ed. *Consciousness*. Atascadero, CA: Ridgeview Publishing.
- Rosenthal, D. M. 1993. "Thinking that one thinks." In M. Davies and G. Humphreys, eds. *Consciousness: Psychological and Philosophical Essays*. Oxford: Blackwell.
- Rosenthal, D. 1994. "First person operationalism and mental taxonomy." *Philosophical Topics*, 22/1-2: 319-50.
- Rosenthal, D. M. 1997. "A theory of consciousness." In N. Block, O. Flanagan, and G. Guzeldere, eds. *The Nature of Consciousness*. Cambridge, MA: MIT Press.
- Russell, B. 1927. *The Analysis of Matter*. London: Kegan Paul.
- Ryle, G. 1949. *The Concept of Mind*. London: Hutchinson and Company.
- Sacks, O. 1985. *The Man who Mistook his Wife for a Hat*. New York: Summit.
- Seager, W. 1995. "Consciousness, information, and panpsychism." *Journal of Consciousness Studies*, 2: 272-88.
- Schacter, D. 1989. "On the relation between memory and consciousness: dissociable interactions and consciousness." In H. Roediger and F. Craik eds. *Varieties of Memory and Consciousness*. Hillsdale, NJ: Erlbaum.
- Schneider W. and Shiffrin, R. 1977. "Controlled and automatic processing: detection, search and attention." *Psychological Review*, 84: 1-64.
- Searle, J. R. 1990. "Consciousness, explanatory inversion and cognitive science." *Behavioral and Brain Sciences*, 13: 585-642.
- Searle, J. 1992. *The Rediscovery of the Mind*. Cambridge, MA: MIT Press.
- Siewert, C. 1998. *The Significance of Consciousness*. Princeton, NJ: Princeton University Press.
- Shallice, T. 1988. *From Neuropsychology to Mental Structure*. Cambridge: Cambridge University Press.
- Shear, J. 1997. *Explaining Consciousness: The Hard Problem*. Cambridge, MA: MIT Press.
- Shoemaker, S. 1975. "Functionalism and qualia," *Philosophical Studies*, 27: 291-315.
- Shoemaker, S. 1981. "Absent qualia are impossible." *Philosophical Review*, 90: 581-99.
- Shoemaker, S. 1982. "The inverted spectrum." *Journal of Philosophy*, 79: 357-381.
- Shoemaker, S. 1990. "Qualities and qualia: what's in the mind," *Philosophy and Phenomenological Research*, Supplement, 50: 109-131.
- Shoemaker, S. 1998. "Two cheers for representationalism," *Philosophy and Phenomenological Research*.
- Silberstein, M. 1998. "Emergence and the mind-body problem." *Journal of Consciousness Studies*, 5: 464-82.
- Silberstein, M. 2001. "Converging on emergence: consciousness, causation and explanation." *Journal of Consciousness Studies*, 8: 61-98.
- Singer, P. 1975. *Animal Liberation*. New York: Avon Books.
- Singer, W. 1999. "Neuronal synchrony: a versatile code for the definition of relations." *Neuron*, 24: 49-65.
- Skinner, B. F. 1953. *Science and Human Behavior*. New York: MacMillan.
- Smart, J. 1959. "Sensations and brain processes." *Philosophical Review*, 68: 141-56.
- Stapp, H. 1993. *Mind, Matter and Quantum Mechanics*. Berlin: Springer Verlag.
- Stoljar, D. 2001. "Two conceptions of the physical." *Philosophy and Phenomenological Research*, 62: 253-81
- Strawson, G. 1994. *Mental Reality*. Cambridge, Mass: MIT Press, Bradford Books.
- Swinburne, R. 1986. *The Evolution of the Soul*. Oxford: Oxford University Press.
- Titchener, E. 1901. *An Outline of Psychology*. New York: Macmillan.
- Travis, C. 2004. "The silence of the senses." *Mind*, 113: 57-94.
- Triesman, A. and Gelade, G. 1980. "A feature integration theory of attention." *Cognitive Psychology*, 12: 97-136.
- Tye, M. 1995. *Ten Problems of Consciousness*. Cambridge, MA: MIT Press.
- Tye, M. 2000. *Consciousness, Color, and Content*. Cambridge, MA: MIT Press.
- Tye, M. 2003. "Blurry images, double vision and other oddities: new troubles for representationalism?" In A. Jokic and Q. Smith eds., *Consciousness: New Philosophical Perspectives*. Oxford: Oxford University Press.
- Van Gulick, R. 1985. "Physicalism and the subjectivity of the mental." *Philosophical Topics*, 13: 51-70.
- Van Gulick, R. 1992. "Nonreductive materialism and intertheoretical constraint." In A. Beckermann, H. Flohr, J. Kim, eds. *Emergence and Reduction*. Berlin and New York: De Gruyter, 157-179.
- Van Gulick, R. 1993. "Understanding the phenomenal mind: Are we all just armadillos?" In M. Davies and G. Humphreys, eds., *Consciousness: Psychological and Philosophical Essays*. Oxford: Blackwell.
- Van Gulick, R. 1994. "Dennett, drafts and phenomenal realism." *Philosophical Topics*, 22/1-2: 443-56.

- Van Gulick, R. 1995. "What would count as explaining consciousness?" In T. Metzinger, ed. *Conscious Experience*. Paderborn: Ferdinand Schöningh.
- Van Gulick, R. 2000. "Inward and upward: reflection, introspection and self-awareness." *Philosophical Topics*, 28: 275-305.
- Van Gulick, R. 2003. "Maps, gaps and traps." In A. Jokic and Q. Smith eds. *Consciousness: New Philosophical Perspectives*. Oxford: Oxford University Press.
- Van Gulick, R. 2004. "Higher-order global states HOGS: an alternative higher-order model of consciousness." In Gennaro, R. ed. *Higher-Order Theories of Consciousness*. Amsterdam and Philadelphia: John Benjamins.
- van Inwagen, P. 1983. *An Essay on Free Will*. Oxford: Oxford University Press.
- Varela, F. and Maturana, H. 1980. *Cognition and Autopoiesis*. Dordrecht: D. Reidel.
- Varela, F. 1995. "Neurophenomenology: A methodological remedy for the hard problem." *Journal of Consciousness Studies*, 3: 330-49.
- Varela, F. and Thomson, E. 2003. "Neural synchronicity and the unity of mind: a neurophenomenological perspective." In Cleermans, A. ed. *The Unity of Consciousness: Binding, Integration, and Dissociation*. Oxford: Oxford University Press
- Velmans, M. 1991. "Is Human information processing conscious?" *Behavioral and Brain Sciences*, 14/4: 651-668
- Velmans, M. 2003. "How could conscious experiences affect brains?" *Journal of Consciousness Studies*, 9: 3-29.
- von Helmholtz, H. 1897/1924. *Treatise on Physiological Optics*. Translated by J. Soothly. New York: Optical Society of America.
- Wilkes, K. V. 1984. "Is consciousness important?" *British Journal for the Philosophy of Science*, 35: 223-43.
- Wilkes, K. V. 1988. "Yishi, duo, us and consciousness." In A. Marcel and E. Bisiach, eds., *Consciousness in Contemporary Science*. Oxford: Oxford University Press.
- Wilkes, K. V. 1995. "Losing consciousness." In T. Metzinger, ed. *Conscious Experience*. Paderborn: Ferdinand Schöningh.
- Watson, J. 1924. *Behaviorism*. New York: W. W. Norton.
- Wegner, D. 2002. *The Illusion of Conscious Will*. Cambridge, MA: MIT Press.
- Wittgenstein, L. 1921/1961. *Tractatus Logico-Philosophicus*. Translated by D. Pears and B. McGuinness. London: Routledge and Kegan Paul.
- Wundt, W. 1897. *Outlines of Psychology*. Leipzig: W. Engleman.
- Yablo, S. 1998. "Concepts and consciousness." *Philosophy and Phenomenological Research*, 59: 455-63.