

BEG

FILOSOFÍA

THOMAS NAGEL

Otras mentes

BEG

FILOSOFÍA

Acercas de este libro, Daniel Dennett escribe: «Nagel es como el último y mejor árbitro es tan bueno como como un extinguidor contra incendios. No hay ningún intento por cuestionar sus resultados, pero en última instancia».

En estos análisis agudos y polémicos, uno de los grandes del momento discute las teorías de Nagel y muchos otros, especialmente en el contexto del emprendido sobre la definición del papel de la ética.

Otras mentes reúne los ensayos críticos y polémicos sobre filosofía cognitiva, ética y filosofía política. Una serie de pensadores desde Aristóteles hasta Wittgenstein, teóricos más significativos del campo de la mente como Robert Nozick y Ronald Dworkin. Otras importantes participaciones de Thomas Nagel en el debate mente-cuerpo, entre los que encontramos a Daniel C. Dennett, John Searle y otros. A lo largo de la obra se expresa su convicción de que las cuestiones intelectuales en nuestros días —desde la fundación científica de la vida a las vicisitudes de la interpretación del derecho— son más filosóficas.

Escritos en un estilo accesible, polémico y vivo, este libro es lo mejor de uno de los pensadores más destacados del siglo XX.

Thomas Nagel es uno de los filósofos contemporáneos más comprometidos. El *Times Literary Supplement* lo calificó como «el más interesante de nuestros días». A lo largo de los últimos treinta años Nagel ha sido una figura central en las polémicas públicas sobre la conciencia, objetividad y ética, liberalismo y razón. Ha publicado numerosas publicaciones de carácter popular, académico y legal, pero su intenso intercambio de ideas y contribuyendo a poner en perspectiva las controversias contemporáneas de interés general con la tradición filosófica.

THOMAS NAGEL

Otras mentes

THOMAS NAGEL

Otras mentes

Ensayos críticos 1969-1994

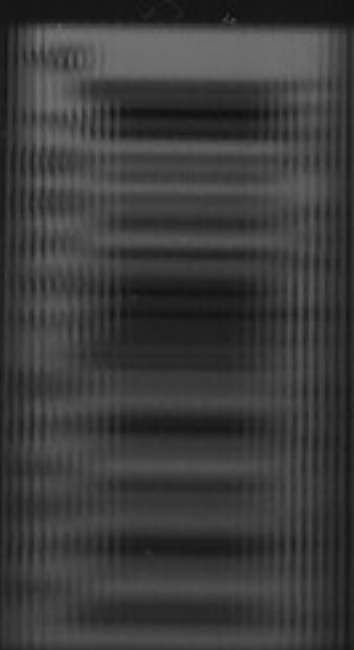
BEG

BIBLIOTECA ECONÓMICA GEDISA

FILOSOFÍA

gedisa
editorial

PVP 9,90€



8

según la cual, si bien los fenómenos mentales son sui generis y no analizables, ni en términos de la conducta ni en alguna otra cosa respecto del cuerpo, están esencialmente conectados con el cuerpo y entre sí. La manera de acercarnos lo más posible a una definición de estos conceptos es descubrir sus conexiones esenciales. Y esto no es lo mismo que reducir los contenidos de la mente a espacios en blanco dentro de la red de circunstancias corporales y manifestaciones, con los cuales se encuentran epistémicamente o causalmente relacionados. Un dolor de cabeza es tan real como un golpe en la cabeza.

Searle: ¿Por qué no somos ordenadores?*

10

Este ensayo fue originalmente una reseña de The Rediscovery of the Mind (M.I.T. Press, 1992), de John Searle. Aunque Searle y yo estamos de acuerdo en muchas cosas, no creo que sea posible distinguir su solución antirreduccionista del dualismo de propiedad. Tampoco creo que pueda ser un hecho bruto de la naturaleza que las propiedades de orden superior del sistema nervioso sean producidas por las particularidades de su operación fisiológica. La relación entre los niveles tiene que ser más "interna" que eso, es decir, una forma de consecuencia inteligiblemente necesaria, antes que una mera correlación. La irreducibilidad de lo ontológicamente subjetivo a lo ontológicamente objetivo sigue siendo un obstáculo a la imaginabilidad de tal conexión.

Según una posición ampliamente aceptada, el cerebro es un ordenador gigante, y la relación entre la mente y el cerebro humanos es semejante a la relación entre el programa del ordenador y el hardware electrónico sobre el que el programa funciona. El filósofo John Searle, un luchador por temperamento, se propuso mostrar que esta posición y la tradición materialista subyacente en ella, son un sinsentido por diferentes razones, algunas de las cuales son obvias y otras más sutiles. Mediante la elaboración de argumentos que él y otros sostuvieron en los últimos veinte años, Searle ataca la mayor parte

* Reimpreso con autorización de *The New York Review of Books*, copyright ©1993 por Nyrev, Inc. Publicado por primera vez el 4 de marzo de 1993.

que todo lo que existe y todo lo que ocurre en el mundo debería poder ser descrito por la ciencia física. Más aún, encuentran evidencia directa de que esto es posible aun para la mente, en la íntima dependencia de lo mental de los procesos neurofisiológicos, sobre los que se ha aprendido mucho desde el siglo XVII hasta hoy. Y encuentran evidencia directa a partir de la exitosa aplicación de la física y la química a otros aspectos de la vida, desde la digestión hasta la herencia. Consecuentemente, la mayoría de los esfuerzos para completar la cosmovisión científica en forma materialista han llevado a cabo una especie de reducción de lo mental a lo físico, donde lo físico, por definición, es lo que puede ser descrito en términos no mentales.

Una reducción es el análisis de algo identificado en un nivel de descripción, en términos de otro nivel de descripción considerado más fundamental, que nos permite decir que el primero en realidad no es sino el segundo: el agua puede describirse como moléculas de H_2O , el calor como movimiento molecular, la luz como radiación electromagnética. Estas son reducciones de lo macroscópico físico a lo microscópico físico, y tienen los siguientes rasgos notables: (1) proveen no sólo información externa sobre las causas o condiciones del fenómeno reducido, sin una explicación interna de lo que realmente es el agua, el calor o la luz; (2) funcionan sólo porque hemos distinguido las apariencias perceptivas de los fenómenos macroscópicos —cómo sentimos el agua y el calor, cómo vemos la luz— de las propiedades que estamos reduciendo. Cuando decimos que el calor consiste en movimiento molecular, queremos decir que el calor, como la propiedad intrínseca de objetos calientes, no es sino el movimiento de sus moléculas. Tales objetos producen en nosotros la sensación de calor cuando los tocamos, pero no identificamos esa sensación con el movimiento molecular; en efecto, la reducción depende de que lo hayamos dejado de lado.

Ahora bien, ¿cómo podríamos reducir los fenómenos mentales a algo descrito totalmente en términos físicos, no mentales? En este caso, no podemos dejar de lado todos los efectos sobre la mente, ya que eso es precisamente lo que tenemos que reducir. Lo que necesitamos para completar la imagen del mundo materialista es algún esquema del tipo: "los fenómenos mentales (pensamientos, sentimientos, sensaciones, deseos, percepciones, etc.) no son sino...", donde tenemos que completar la

frase con una descripción que sea explícitamente física, o que utilice sólo términos aplicables a algo totalmente físico.² Los diferentes intentos de llevar a cabo esta tarea, en apariencia imposible, constituyen la historia de la filosofía de la mente de los últimos cincuenta años.

La participación de Searle en esta empresa parte del conductismo, la concepción de que los conceptos mentales no se refieren a nada dentro de nosotros, y que cada tipo de estado mental puede ser identificado con una disposición del organismo para conducirse de determinadas maneras bajo ciertas condiciones físicas. Cuando esta posición parecía convertirse en un franco rechazo a la existencia de la mente, algunos filósofos propusieron teorías de la *identidad*, según las cuales los procesos mentales son idénticos a los procesos cerebrales, del mismo modo que la luz es idéntica a la radiación electromagnética. Pero las teorías de la identidad tenían el problema de explicar en términos no mentales, qué quiere decir de un proceso cerebral determinado, que es un pensamiento o una sensación. Después de todo, no quiere decir solamente que es cierto tipo de proceso neurofisiológico. Dado el objetivo de estas teorías, no podía querer decir que el proceso cerebral tiene algún efecto mental. La solución propuesta era un renacimiento del conductismo, con una nueva forma: la sed, por ejemplo, se identificaba no sólo con una disposición a beber, sino con un estado cerebral, pero ese estado cerebral determinado, al ser idéntico a la sed, ahora consistía simplemente en el hecho de que era típicamente causado por deshidratación, y que causaba típicamente la disposición a beber. De este modo, se pensaba que la identificación de estados mentales con estados cerebrales podía evitar toda referencia a rasgos no físicos.

A estos análisis "conductistas causales" les sucedió, eventualmente, una teoría más técnica, llamada "funcionalismo", según la cual los conceptos mentales no pueden estar unidos individualmente a la conducta y circunstancias, sino como una parte de una única red interconectada. La conducta causada por la sed, por ejemplo, depende del resto de la condición mental de la persona, es decir, sus creencias sobre dónde encontrar agua, y sobre si el agua es potable, la fuerza de sus deseos de vivir o morir, etc. Cada estado mental forma parte de un sistema integrado, que controla la interacción del organismo con el

medio: sólo mediante el análisis del papel que desempeñan los estados como sed, dolor, otros tipos de sensación, creencia, emoción y deseo, dentro del sistema en su totalidad, se puede describir adecuadamente su conexión con la conducta y las circunstancias externas. Un sistema de este tipo permite identificar estados mentales con estados cerebrales, siempre que los últimos tengan roles causales o funcionales del tipo especificado por la teoría (que debe ser construida) de cómo funciona el sistema integrado. Finalmente, el funcionalismo condujo a lo que Searle denomina IA Fuerte (Inteligencia Artificial Fuerte), a saber, la identificación de estados mentales con estados computacionales de un programa de ordenador que controla la conducta del organismo, un programa que se lleva a cabo físicamente en el hardware (o wetware) del cerebro.³

Todas estas teorías intentan reducir la mente a algún aspecto de un mundo que pueda describirse completamente por la física, es decir, el mundo de las partículas y campos. Estas teorías no fueron desarrolladas en detalle, sólo son esperanzadas descripciones de lo que debería ser una teoría de la mente, junto con algunos ejemplos extremadamente esquemáticos. En tanto que cada nueva propuesta fue criticada por defensores de algún enfoque reduccionista alternativo, Searle argumenta que todas ellas tienen un grave defecto: todas dejan a la conciencia de lado.

II

Ninguna teoría que deje la conciencia de lado puede pretender ser una teoría de la mente, y no es posible hacer un análisis de la conciencia en términos no mentales. Consecuentemente, ninguna reducción materialista de lo mental puede tener éxito. Searle arguye que ninguna de estas teorías podría decir lo que realmente es dolor, hambre, creencia, visión, etc., porque todo lo que tratan es externamente observable —la conducta del organismo y su estructura causal—, y una descripción exclusivamente en esos términos no puede garantizar que tiene conciencia: cada una de estas teorías conductistas, funcionalistas o computacionales podrían satisfacerse por un autómata inconsciente dotado de la suficiente complejidad física.

La pregunta crucial no es “¿Bajo qué condiciones atribuiríamos estados mentales a otras personas?, sino, “¿Qué es lo que *realmente* tiene la gente cuando tiene estados mentales?”, “¿Qué son los fenómenos mentales?”, en tanto distinta de “¿Cómo conocerlos, y cómo funcionan causalmente en la vida del organismo?”. (p.23)

Les atribuimos conciencia a otras personas y a los animales, sobre la base de su conducta, lo cual es simplemente una evidencia de conciencia, más bien que una prueba de conciencia, y debe ser complementada con evidencia de similaridad fisiológica: ya que creemos en la uniformidad de la naturaleza, inferimos naturalmente que criaturas que se conducen en forma similar a nosotros, y que tienen órganos sensoriales y sistemas nerviosos similares físicamente a los nuestros, también tienen experiencias conscientes de alguna clase. Pero, Searle argumenta, ninguna cantidad de hechos sobre la conducta física o la organización funcional implica por sí misma que un sistema sea consciente. Y, una teoría que sostiene, por ejemplo, que la visión no es “nada más que” un estado del organismo, a fin de ser adecuada, debe tener la consecuencia de que, si el organismo está en ese estado, *no puede* dejar de ser consciente. De lo contrario, dejará de lado lo más importante sobre la visión y, por más ventajas que ofrezca, no podrá representar un enfoque válido sobre lo que es la visión.

Los reduccionismos materialistas no sólo no pueden implicar que el sistema es consciente; es evidente de antemano que ningún desarrollo posterior en esta misma dirección, ni ninguna complicación adicional estructural o conductista, podrían hacerlo. La razón es que hay una diferencia crucial entre fenómenos conscientes y fenómenos de conducta o fisiológicos, que hace que los primeros sean irreducibles a los segundos: la conciencia es, en términos de Searle, “ontológicamente subjetiva”. Es decir, sus rasgos esenciales no pueden ser descritos totalmente desde un punto de vista externo de una tercera persona. Incluso la descripción fisiológica de lo que ocurre dentro del cráneo es, en este sentido, externa: la describimos desde afuera. No basta con resumir las observaciones, conductuales o fisiológicas, de una tercera persona, que nos permiten atribuirles estados mentales conscientes a los demás. El punto de vista de la primera persona, que nos revela cómo es un estado mental consciente para el sujeto, es indispensable.

Esto se torna evidente cuando preguntamos: ¿qué es la conciencia? Aunque podamos describir ciertos rasgos e identificar más tipos específicos de fenómenos mentales, en tanto sus instancias, es tan básica que no puede ser definida en términos de otra cosa. Usted, lector, es consciente, en este preciso momento, y su condición de conciencia incluye cosas como su manera de percibir esta página, cómo siente el papel entre sus dedos, cómo siente su camisa en el cuerpo y la silla sobre la que está sentado, los sonidos que está escuchando, la música, o las sirenas de la policía en el fondo; y su experiencia de leer esta oración. La posición de Searle es que el análisis de la tercera persona, sea conductual, causal o funcional, no importa cuán extenso sea, no puede decirnos qué son estas experiencias en sí mismas, en qué consisten, en tanto distintas de sus causas y efectos. Esto es absolutamente obvio, porque los hechos subjetivos sobre lo que es para una persona estar en ciertas condiciones no pueden ser idénticos a los hechos sobre lo que las cosas son, ni desde el punto de vista de una persona, ni para nadie, sino en sí mismos. Los hechos sobre nuestra conducta externa o la actividad eléctrica u organización funcional de nuestro cerebro pueden estar conectados de cerca con nuestra experiencia consciente, pero no son hechos sobre lo que es para mí escuchar una sirena de policía.

Searle cree que la insistencia en el reduccionismo materialista, a pesar de su evidente falsedad, requiere una explicación. El la compara con la repetición constante de una conducta destructiva por parte de un neurótico compulsivo, y espera, al traer a la luz sus causas subyacentes, poder romper con esa compulsión. Es evidente, tanto por lo que dicen como por lo que hacen, que los reduccionistas están convencidos de antemano de que alguna teoría materialista debe ser correcta: lo que tienen que hacer es encontrarla. Esta suposición es parte de una cosmovisión científica que, según ellos, no tiene alterna-tiva. Pero, según Searle, hay dos errores conceptuales cruciales que subyacen en esta suposición. El primero de ellos es que debemos elegir entre materialismo y dualismo:

No quiero dejar de insistir en que uno puede aceptar los hechos obvios de la física —por ejemplo, que el mundo consiste enteramente en partículas físicas, en campos de fuerza— sin negar al mismo tiempo los hechos obvios de nuestras propias experiencias, por ejemplo, que somos todos conscientes y que nuestros estados

conscientes tienen propiedades fenomenológicas *irreducibles* [es decir, subjetivas] bastante específicas. El error consiste en suponer que estas dos tesis son incompatibles, y que el error proviene de aceptar los supuestos del vocabulario tradicional. Mi posición no es, de ningún modo, una forma de dualismo. Rechazo tanto el dualismo de propiedad como el sustantivo. Pero, precisamente por las razones por las que rechazo el dualismo, rechazo el monismo y también el materialismo. [El error profundo es suponer que debemos optar entre estas posiciones . . .]

Una vez que aceptamos nuestra cosmovisión, el único obstáculo para garantizarle a la conciencia su carácter de rasgo biológico de un organismo, es la suposición, pasada de moda, dualista/materialista, de que el carácter "mental" de la conciencia le impide ser una propiedad "física". (pp. 28, 91)

Esta tesis radical de que la conciencia es una propiedad física del cerebro, a pesar de su subjetividad, y de que es irreducible a cualquier *otra* propiedad física, es el corazón metafísico de la posición de Searle. Sin embargo, lo importante es que Searle sostiene que los materialistas se dejan conducir hacia formas imposibles de reducción psicofísica, porque suponen que, si los estados mentales no pueden explicarse en tales términos, entonces la alternativa ineludible es el dualismo: en consecuencia, deberían admitir que las sustancias o propiedades no físicas son rasgos básicos de la realidad. Y el temor al dualismo, con sus asociaciones religiosas, espiritualistas y no científicas, los lleva a adoptar el materialismo reduccionista a cualquier precio: "Entonces, el materialismo es, en cierto sentido, la flor más fina del dualismo".

Para escapar de este brete, según Searle, tenemos que liberarnos de la urgencia de preguntarnos si hay una o dos clases últimas de cosas y propiedades. En primer lugar, no deberíamos empezar por contarlas.

Searle está absolutamente en lo cierto sobre el temor al dualismo (en efecto, creo que él mismo no es inmune a él). Su expresión más extraña es, sin embargo, otra teoría llamada "materialismo eliminativo". Se trata de la posición de que, dado que los estados mentales no pueden acomodarse dentro del mundo descrito por la física, no existen, así como no existen ni las brujas ni los fantasmas. Pueden ser rechazados como postulados de una teoría primitiva, a la que habitualmente nos referimos como "folk psychology". Sobre esta teoría, P. F. Straw-

son ha dicho que es la provincia de "gente tan sencilla como Shakespeare, Tolstoy, Proust y Henry James".⁵ Pacientemente, Searle pulveriza esta posición, pero el núcleo de su argumentación es que toda la tradición materialista es eliminativa: todas las teorías materialistas niegan la realidad de la mente, pero la mayoría de ellas disfraza el hecho (por sí mismos o por los demás), al identificar la mente con alguna otra cosa.

El segundo error conceptual crucial que trasciende la búsqueda compulsiva de teorías materialistas, según Searle, es el error, simple pero enormemente destructivo, sobre la objetividad:

Hay una constante confusión entre la afirmación de que deberíamos tratar, en lo posible, de eliminar los prejuicios subjetivos personales en la búsqueda de la verdad, y la afirmación de que el mundo real no contiene ningún elemento que sea irreductiblemente subjetivo. Esta confusión, a su vez, se basa en una confusión entre el sentido epistémico y el sentido ontológico de la distinción subjetivo/objetivo. En el sentido epistémico, la distinción marca diferentes grados de independencia de afirmaciones, desde los caprichos de valores especiales, prejuicios personales, puntos de vista y emociones. En el sentido ontológico, la distinción marca diferentes categorías de realidad empírica. (p. 19)

Esto me parece totalmente convincente, y muy importante. Naturalmente, la ciencia debe buscar objetividad epistémica —conocimiento objetivo— usando métodos que compensen los diferentes puntos de vista y que permitan a diferentes observadores llegar a la misma concepción de algo. Pero es una gran confusión concluir, a partir de eso, que nada que tenga o incluya un punto de vista pueda ser objeto de investigación científica. Los puntos de vista subjetivos son, ellos mismos, partes del mundo real y, si ellos y sus propiedades deben ser descritos adecuadamente, su carácter ontológicamente subjetivo —la subjetividad de su *naturaleza*— debe ser reconocida. Más aún, es posible hacerlo objetivamente en un sentido epistémico: aun que sólo usted esté teniendo la experiencia de mirar la página ante sus ojos, otros pueden saberlo y pueden saber bastante bien cómo es esa experiencia para usted. Es una verdad objetiva que usted está teniendo cierta experiencia visual subjetiva en este momento.

Si aceptamos esta distinción, la pregunta es: ¿cómo podemos formarnos una concepción científica objetiva, desde el punto de vista epistémico, de un mundo que no contiene solamente los hechos objetivos ontológicamente familiares, descritos por la física, química, biología, sino también los hechos ontológicamente subjetivos de la conciencia? Y esta pregunta nos trae, finalmente, a la posición de Searle, que denomina "naturalismo biológico", y que combina la aceptación de la irreductible subjetividad de lo mental con el rechazo a la dicotomía entre lo mental y lo físico: "la conciencia... es un rasgo biológico del cerebro humano y de los de ciertos animales. Está causada por procesos neurobiológicos y es tanto parte del orden biológico natural como cualquier otro rasgo biológico, ya sea fotosíntesis, digestión o mitosis". (p. 90)

Y, a pesar de su antirreduccionismo, también dice lo siguiente:

La conciencia es una propiedad de nivel superior, o emergente, del cerebro, en el sentido totalmente inocuo de "nivel superior" o "emergente", en el cual la solidez es una propiedad emergente de nivel superior de moléculas de H_2O , cuando se encuentran en una estructura entrelazada (hielo), y la liquidez es, análogamente, una propiedad emergente de nivel superior de moléculas de H_2O cuando están, por así decirlo, rodando unas con otras (agua). La conciencia es una propiedad mental, por lo tanto física, del cerebro, en el sentido en que la liquidez es una propiedad de sistemas moleculares. (p.14)

Si pudiéramos aclarar esta posición, como para distinguirla de las alternativas, debería representar un aporte fundamental a las respuestas posibles al problema mente-cuerpo. Pero no creo que lo sea.

Supongamos que garantizamos que los estados de la conciencia son propiedades del cerebro causadas por, pero no reducibles a, su actividad neuronal. Esto quiere decir que mi cerebro, por ejemplo, tiene un punto de vista, del que mis experiencias actuales son aspectos. Pero, ¿cuál es la justificación para llamar *físicos* a estos rasgos del cerebro, que son irreducibles y subjetivos? ¿Qué significa llamarlos físicos? Por cierto, son de "nivel superior" en el sentido de que pueden ser atribuidos sólo al sistema como un todo y no a sus partes microscópicas; también son "emergentes" en el sentido de que se explican

sólo por las interacciones causales de aquellas partes. Pero, por mayor que sea la variedad de fenómenos físicos, la subjetividad ontológica es una de sus condiciones definitorias centrales. Asimismo, como ya vimos, Searle insiste en que la conciencia es ontológicamente subjetiva.

Searle no dice demasiado sobre este punto. Quizá cree que, si los cerebros están constituidos por partículas físicas, se deduce automáticamente que todas sus propiedades son físicas. Y cita una observación de Noam Chomsky, quien sostiene que, en cuanto comprendemos algo, lo llamamos "físico". Pero, si "físico" es, en este sentido, un término meramente honorífico (este es otro término que Chomsky ha usado), entonces, ¿cuál es el contenido metafísico de la afirmación de Searle de que las propiedades mentales son físicas, y su enfático rechazo al dualismo de propiedad? El dice, después de todo, que la distinción ontológica entre subjetivo y objetivo señala "diferentes categorías de la realidad empírica". Pero decir, además, que nos "quedamos con un universo que contiene un componente físico irreductiblemente subjetivo como componente de la realidad física", sólo conlleva una postura esencialmente dualista en el lenguaje, que expresa una fuerte aversión al dualismo.⁶

Quizá deberíamos adoptar el uso de Searle de la palabra "físico", pero el problema central es más que un problema verbal. Es el problema de cómo construir una cosmovisión científica inteligible y completa, una vez que rechazamos la reducibilidad de lo mental a lo no mental. Tal como señala Searle, para lograr este objetivo, no podemos seguir por el camino que la ciencia ha seguido desde el siglo XVII, porque este dependía de la exclusión de la mente del observador del mundo que se observa y describe. Proponer a la conciencia como una propiedad subjetiva intrínseca del cerebro, causada por su actividad neuronal, es el primer paso en un camino diferente: el camino correcto, en mi opinión. Pero hay grandes problemas a resolver, y no se trata sólo de problemas empíricos, sino también filosóficos.

Aun si aprendemos mucho más de lo que sabemos hoy en día sobre las causas fisiológicas de la conciencia, no hará, como Searle ya sabe, a la relación de la conciencia con la conducta de neuronas análoga a la relación de la liquidez con la conducta de moléculas de H₂O. En este último caso, la relación es transparente: podemos *ver* cómo la liquidez es el resultado lógico de las moléculas "rodando unas con otras" en el nivel microscópi-

co. No podemos esperar nada comparable en el caso de las neuronas, aun cuando sea empíricamente evidente que los estados de conciencia son las consecuencias necesarias de la actividad neuronal. Searle hace un análisis interesante sobre esta diferencia que, dice, resulta no sólo de una limitación de nuestros poderes para concebirlo: podemos representarnos la relación necesaria entre los niveles macro y micro del agua, ya que nos los imaginamos desde afuera, pero no podemos hacerlo con la subjetividad, que nos tenemos que imaginar desde adentro, sea la nuestra o la de los demás. Estoy de acuerdo, pero creo que esto significa que no entendemos realmente la afirmación de que los estados mentales son estados del cerebro: todavía somos incapaces de concebir *cómo* la conciencia surge en la materia, incluso si estamos seguros de que lo hace.⁷

III

La segunda serie de argumentos de Searle en contra del modelo computacional de la mente depende de la naturaleza específica de los ordenadores, y es más típicamente propia de Searle: surge de su antigua preocupación por la teoría del significado y la "intencionalidad" de los estados mentales: es decir, su capacidad para significar algo o referirse a algo fuera de ellos mismos, y su consecuente susceptibilidad respecto de los juicios de verdad o falsedad, y corrección o incorrección.⁸

¿Cómo pueden los ordenadores responder a las preguntas que les hacemos, por ejemplo, las preguntas aritméticas? La explicación tiene dos partes. En primer lugar, es posible formular cada una de las preguntas usando una secuencia de símbolos —letras o numerales— que se seleccionan de una corta lista y se distinguen por sus formas, y diseñar una serie finita de reglas para manipular esos símbolos, la cual tiene la siguiente propiedad: si comenzamos con la secuencia que corresponde a la pregunta, y seguimos las reglas para mover, eliminar y agregar símbolos, llegaremos, después de una serie finita de pasos determinados uno tras otro, a un punto en el que las reglas nos dirán que debemos detenernos; la última secuencia que produjimos corresponderá a la respuesta,⁹ que simplemente debemos leer. Pero, a fin de seguir las reglas para manipular los símbo-

los, no necesitamos saber qué significan, si es que significan algo; simplemente, debemos identificar sus formas.

Estas reglas se llaman reglas de la *sintaxis* (en oposición a las reglas de la *semántica*, que necesitamos para interpretar una secuencia que signifique algo). Lo bueno de esto es que podríamos entrenar a alguien en división, por ejemplo, en forma totalmente mecánica, usando una clase de reglas y tablas, sin que él sepa que los símbolos que está escribiendo representen números o alguna otra cosa: si esa persona sigue las reglas sintácticas, llegará a lo que nosotros leeríamos (aunque él no) como la respuesta.¹⁰

La segunda parte de la explicación es que hay, además de la escritura sobre papel, diferentes maneras de codificar la aritmética en un conjunto de símbolos y reglas sintácticas, y, para alguna de esas maneras, es posible diseñar máquinas físicas a fin de que puedan llevar a cabo mecánicamente todos los pasos en la manipulación de los símbolos que prescriben las reglas. En lugar de la persona que sigue las reglas sintácticas mecánicamente, sin saber lo que significan los símbolos, un mecanismo físico puede hacer las mismas operaciones de los símbolos en forma automática. Pero, nótese que este mecanicismo no sólo no sabe lo que los símbolos significan, sino que ni siquiera sabe, a diferencia del escriba de nuestro ejemplo, que ignora la semántica, que está siguiendo reglas para su manipulación, reglas sintácticas. No sabe nada. En realidad, no está siguiendo reglas, sino que está funcionando solamente de acuerdo con las leyes de la física, en las formas diseñadas por ingenieros expertos, que nos permiten interpretar los resultados tanto sintáctica como semánticamente.

El conocido argumento de Searle de la “habitación china” describe una persona consciente que, sin saber chino, sigue reglas de manipulación de los caracteres chinos y arma operaciones inteligibles para personas que saben chino. El punto crucial en contra del modelo computacional de la mente es que la sintaxis sola no puede llegar a la semántica.¹¹ En *The Rediscovery of the Mind*, Searle extiende el argumento para mostrar que la física sola no puede llegar a la sintaxis. Seguir reglas, aunque sean reglas puramente sintácticas, es un proceso mental irreducible, un proceso “intencional”, en el que el significado de las reglas en sí mismas debe ser captado por una mente consciente. No se trata sólo de una cuestión de regularidad en la conducta física.

Un estado físico de un sistema es un estado computacional sólo cuando se le asigna a ese estado algún papel, función o interpretación computacional . . . nociones tales como *computación*, *algoritmo* y *programa* no nombran rasgos físicos intrínsecos de sistemas: no descubrimos estados computacionales dentro de la física, se los asignamos a la física . . .

El objetivo de la ciencia natural es descubrir y caracterizar rasgos que son intrínsecos del mundo natural. Por la misma definición de computación y conocimiento, no hay ninguna forma en que la ciencia cognitiva computacional pueda ser una ciencia natural, porque la computación no es un rasgo intrínseco del mundo. Es asignado en forma relativa a los observadores. (pp. 210, 212; bastardilla en el original)

La distinción de Searle entre lo que es intrínseco a la cosa observada y lo que es relativo a un observador o intérprete es fundamental. El argumento que la intencionalidad intrínseca —es decir, la capacidad de captar el significado de enunciados, y seguir reglas conscientemente— ocurre sólo en la mente. Podemos decir que palabras en una página o resistencias eléctricas en un chip de ordenador significan algo o que obedecen reglas de gramática o de aritmética, sólo en el sentido derivativo de que nuestra mente puede interpretarlos de ese modo, en virtud de su disposición. Esto significa que la afirmación de que el cerebro es un ordenador implica que el cerebro tiene intencionalidad y sigue reglas de computación que no son intrínsecas, sino sólo relativas a la interpretación de su usuario. ¿Y quién se supone que es ese usuario? Si el cerebro es un ordenador, no tiene intencionalidad intrínseca. Si la tiene, deber ser algo más que un ordenador. Searle elige esta segunda alternativa. También discute que las teorías que tratan de representar la mente como un ordenador, subrepticamente suponen una mente o un “homúnculo” que la interprete.

Hay mucho más todavía en este argumento y, aunque su conclusión negativa me parece convincente, se podrían plantear algunas preguntas que no he tratado aquí con respecto a la teoría positiva de Searle de la intencionalidad intrínseca, que, supuestamente, es compatible con su naturalismo biológico. Al igual que la conciencia, es muy difícil ver cómo la intencionalidad intrínseca podría ser una propiedad del organismo físico. Pero, en lugar de ahondar en estas cuestiones, consideraré la posición de Searle sobre el inconsciente y su relación con la conciencia, ya que ellos unen ambas partes del argumento.

Searle le ha dado gran importancia a la afirmación de que la conciencia subjetiva no es reductible a ninguna otra cosa. Pero la mayoría de nuestros estados mentales no son conscientes. Tomemos todas las creencias e intenciones que tenemos, pero en las que no estamos pensando en este momento. Por ejemplo, la creencia de que hay una torre inclinada en Pisa. Recién se ha vuelto consciente, pero probablemente tenemos esta creencia desde hace años. Si tales creencias pueden existir en forma inconsciente, entonces la conciencia no es un rasgo esencial de la vida mental, y debe ser posible que los estados mentales estén incorporados en un cerebro puramente material. De modo que se podría decir que, incluso para aquellos estados mentales que son conscientes, su carácter subjetivo, experimentado, no es esencial para su intencionalidad. Quizá la conciencia sea sólo un tipo de "tinte" subjetivo que algunas veces se les agrega al blanco y negro verdaderamente funcional de los estados mentales, de modo que una teoría de la mente podría dejar de lado a la conciencia, en tanto algo no esencial.

Esta es la respuesta de Searle a esta sugerencia:

Entendemos la noción de estado mental inconsciente solamente como un contenido posible de la conciencia, sólo como el tipo de cosa que, aunque no es consciente y es quizás imposible de traerse a la conciencia, por varias razones [tales como la represión] es, sin embargo, la *clase de cosa* que puede ser o podría haber sido consciente. (p. 156)

Esto es lo que Searle llama el "principio de conexión"; sus argumentos a favor de él es que incluso los estados mentales inconscientes deben tener un carácter distintivamente subjetivo, que puedan contar como creencias, pensamientos, deseos, etc. Para poder ser sobre algo, y por lo tanto, ser verdadero o falso, correcto o incorrecto, un estado debe pertenecer, al menos potencialmente, al punto de vista de algún sujeto. Searle reconoce que era un estado neurofisiológico de mi cerebro lo que hizo verdadero, hace dos horas, que yo creyera que había una torre inclinada en Pisa, pero argumenta que la neurofisiología sola no puede elevar ese estado al rango de creencia, porque ninguna descripción fisiológica por sí misma implica que el estado cerebral tenga alguna intencionalidad o significado, aun si le agregamos a la descripción la conducta física que podría causar. Su conclusión es que se les puede atribuir inten-

cionalidad a los estados neurofisiológicos sólo en forma derivativa, en virtud de los estados conscientes que pueden producir. "La ontología del inconsciente consiste en rasgos objetivos del cerebro, capaces de causar pensamientos conscientes subjetivos". Esto último tiene la sorprendente consecuencia de que un mecanismo profundo, supuestamente psicológico, como el Instrumento de adquisición del lenguaje de Chomsky, que le permite al niño aprender la gramática de un lenguaje a partir de los segmentos de habla con los que se encuentra en una edad temprana, no es una clase de reglas mentales inconscientes, sino simplemente un mecanismo físico, porque es incapaz de dar origen a un pensamiento consciente subjetivo, cuyo contenido consiste en esas mismas reglas. De manera que, en la posición de Searle, la conformidad del niño a las reglas, al aprender el lenguaje, no es un ejemplo de intencionalidad intrínseca, sino sólo una intencionalidad que le atribuye el observador-lingüista.

En suma, la conciencia es la esencia de lo mental, incluso si la mayoría de los estados mentales no son conscientes en algún momento determinado. No podemos estudiar ningún aspecto de la experiencia mental sin incluirla, o ella o a su posibilidad, en la definición de lo que estamos tratando de entender. En particular, la intencionalidad es inseparable de ella.

The Rediscovery of the Mind es incisivo, agresivo y bellamente claro, en el mejor estilo de Searle de "¿Qué es todo este sinsentido?". Es un valioso antídoto contra una de las ilusiones predominantes de nuestro tiempo.

Notas

1. Véase Geoffrey Madell, *Mind and Materialism* (Edinburgh University Press, 1988).
2. Otra estrategia reduccionista, que no puedo analizar aquí en extenso, consiste en sustituir una teoría de lo que son los estados mentales, por una teoría de las bases observables externamente por las que *atribuimos* estados mentales a las personas, y sostener que este sistema de "condiciones de asertividad" es todo el análisis que estos conceptos requirieren. No *identificamos* los fenómenos mentales con algo físico, porque no necesitamos identificarlos con nada. Pero las condiciones de aplicación de los conceptos mentales son, según esta posición, compatibles con el hecho de que el mundo es nada más que un sistema material. Esta es fundamentalmente la estrategia de Daniel Dennett en *Consciousness Explained* (Little, Brown, 1991) [analizada en el cap. 8 de este volumen].

II

ÉTICA Y FILOSOFÍA POLÍTICA

3. El conductismo está más o menos representado por Gilbert Ryle, la teoría de la identidad por J. J. C. Smart, y el funcionalismo por Hilary Putnam, pero hay muchos autores, y la bibliografía es muy extensa. Véanse algunas de las excelentes colecciones recientes sobre la filosofía de la mente: Ned Block (comp.), *Readings in Philosophy of Psychology*, 2 vols. (Harvard University Press, 1980); W. G. Lycan (comp.), *Mind and Cognition*, (Blackwell, 1990); David Rosenthal (comp.), *The Nature of Mind* (Oxford University Press, 1991). Putnam ha abandonado ahora el funcionalismo, véase *Representation and Reality* (M.I.T. Press, 1988).

4. Una observación: yo soy uno de los que han sido citados como defensores de esta línea de argumentación. Los otros son Saul Kripke y Frank Jackson.

5. El materialismo eliminativo fue propuesto por primera vez por Paul Feyerabend y Richard Rorty. Más versiones de esta posición son defendidas por Steven Stich, *From Folk Psychology to Cognitive Science: The Case against Belief* (M.I.T. Press, 1983); Paul M. Churchland, *Matter and Consciousness: A Contemporary Introduction to the Philosophy of Mind* (M.I.T. Press, 1984) y Patricia S. Churchland, *Neurophilosophy* (M.I.T. Press, 1986). El comentario de Strawson se encuentra en *Skepticism and Naturalism: Some Varieties* (Columbia University Press, 1985) p. 56.

6. Searle me identifica como un defensor del dualismo de propiedad. Yo prefiero el término "teoría del doble aspecto" para expresar la posición, que proviene de Spinoza, de que los fenómenos mentales son los aspectos subjetivos de estados que pueden ser también descritos físicamente. Pero todo lo que yo sostendría a favor de esta idea es que es algo más aceptable que las otras teorías habituales inaceptables. Comparto la aversión de Searle tanto hacia el dualismo como el materialismo, y no creo que haya ninguna solución al problema mente-cuerpo a la vista.

7. Colin McGinn, en *The Problem of Consciousness* (Blackwell, 1991) sugiere que somos constitucionalmente incapaces de llegar a ese conocimiento. Aunque él pueda estar en lo cierto, creo que su pesimismo es prematuro.

8. Véase el ensayo de Searle, "Minds, Brains, and Programs", en *Behavioral and Brain Sciences* 3, Nº 3 (septiembre de 1980), pp. 417-24. Véase también *Intentionality: An Essay in the Philosophy of Mind* (Cambridge University Press, 1983).

9. Este procedimiento se denomina "algoritmo".

10. La teoría general de este tipo de computabilidad fue desarrollada por Alan Turing.

11. Véase su "Minds, Brains, and Programs".