
LA TEORÍA DE LA MENTE Y EL LEGADO DE DARWIN

JOHN SEARLE

ABSTRACT. We do not have an adequate theory of consciousness. Both dualism and materialism are mistaken because they deny consciousness is part of the physical world. False claims include (i) behaviorism, (ii) computationalism, (iii) epiphenomenalism, (iv) the readiness potential, (v) subjectivity, and (vi) materialism. Ontological subjectivity does not preclude epistemic objectivity. Observer relative phenomena are created by consciousness, but consciousness is not itself observer relative. Consciousness consists of feeling, sentience, or awareness with (i) qualitativity, (ii) ontological subjectivity, (iii) unified conscious field, (iv) intentionality, and (v) intentional causation. All conscious states are caused by lower level neurobiological processes in the brain, and they are realized in the brain as higher level features. Efforts to get a detailed scientific account of how brain processes cause consciousness are disappointing. The Darwinian revolution gave us a new form of explanation; two levels were substituted: a causal level, where we specify the mechanism by which the phenotype functions, and a functional level, where we specify the selectional advantage that the phenotype provides. Sociobiology attempted to explain general features of human society, ethics, etc. It failed. For the incest taboo, it confuses inhibition with prohibition. It did not explain the moral force of the taboo. To explain the function of consciousness we cannot ask, "What would be subtracted if we subtracted consciousness but left everything else the same?" We cannot leave everything else the same because consciousness is necessary for higher functions of human and animal life. The unified conscious field gives the organism vastly increased power.

KEYWORDS. Objective/subjective, unconscious.

Tengo dos objetivos en este artículo. En primer lugar, quiero decir algo sobre nuestro entendimiento actual de la naturaleza de la mente, particularmente de la conciencia. En segundo lugar, quiero relacionar mi descripción de la mente con la evolución darwiniana. Muchas de las ideas de este artículo han aparecido de forma preliminar en varios de mis libros y artículos, especialmente importantes son Searle 1996, 1992 y 2005.

EL "ESCÁNDALO" DE LA CONCIENCIA

La conciencia es un escándalo porque no tenemos una teoría neurobiológica adecuada de ésta; en lugar de ello hay un gran número de afirmaciones falsas al respecto. Comenzaré con media docena de afirmaciones falsas

Departamento de Filosofía, Universidad de California, Berkeley, CA 94720. / searle@berkeley.edu. / JS escribió el artículo.

Ludus Vitalis, vol. XXI, num. 40, 2013, pp. 11-27.

sobre la conciencia que he encontrado en mi vida intelectual, muchas de las cuales han sido ampliamente conocidas y extremadamente influyentes.

Conductismo. Según la reconstrucción tradicional, la conciencia realmente no existe, sino que sólo hay disposiciones conductuales, además de comportamientos humanos y animales (Skinner 1992).

Computacionalismo (inteligencia artificial fuerte). La conciencia como tal no existe, sino que es un programa o un conjunto de programas computacionales que operan en el cerebro (Minsky y Papert 1987).

Epifenomenalismo. La conciencia existe, pero no puede tener ningún efecto real en el mundo porque no es un fenómeno físico, y por ello no puede afectar al mundo físico. Debe ser un epifenómeno porque el mundo físico está “causalmente cerrado”. Desde un punto de vista evolutivo, no tiene ninguna función (Chalmers 1996).

El potencial para la acción. La conciencia existe, pero tiene muy poca importancia porque la investigación sobre el potencial para la acción en la corteza motora suplementaria muestra que nuestras acciones se inician antes de que seamos conscientes de lo que estamos haciendo. El cerebro decide llevar a cabo una acción antes de que la mente sea consciente de ello (Libet, et al. 1983).

Objetividad y subjetividad. La conciencia no es un tema apropiado para la investigación científica seria, y es mejor dejársela a los teólogos y los filósofos. La razón es que por definición la ciencia es objetiva, y por definición la conciencia es subjetiva, por lo tanto, no puede haber una ciencia de la conciencia. Esta visión es parte de la tradición oral. Cuando empecé a interesarme en la neurobiología de la conciencia, discutí el asunto con varios neurobiólogos, tanto en Europa como en Estados Unidos, muchos de los cuales me aseguraron que la conciencia no es un tema apropiado para la investigación científica seria debido a su subjetividad.

Materialismo. Si la conciencia es real, realmente debe ser algo más, porque el inventario final de la ontología básica del mundo —un inventario que incluye a las partículas subatómicas, la gravedad, el electromagnetismo, las fuerzas nucleares fuertes y débiles, y otras características fundamentales de la realidad— es completamente material y por lo tanto no incluye a la conciencia (Dennett 1991).

Argumentaré que la conciencia es un fenómeno biológico real en el mismo sentido que lo son la fotosíntesis o la digestión, y que se puede demostrar que estas seis posturas son más o menos falsas. Ahora bien, esto plantea una pregunta interesante: ¿por qué hay tanta confusión respecto a la conciencia?

¿Por qué es tan difícil que la gente acepte una acepción naturalista? Somos víctimas de dos tradiciones que parecen ser incompatibles entre sí,

cuando en realidad se complementan. Una es la tradición de Dios, el alma y la inmortalidad, que dice que la conciencia no es parte del mundo natural. La conciencia ni siquiera es una propiedad del cuerpo o del cerebro. En la visión cartesiana, es una propiedad del alma y definitivamente el alma no es parte del mundo natural. Esta es la tesis por la cual, por cierto, Descartes fue forzado a admitir que los animales no son conscientes porque todo lo que el animal tiene es un cuerpo. Para Descartes nuestros cuerpos no son conscientes, nuestro cerebro no es consciente; sólo el alma es consciente. Usted podría estar pensando que esta tradición está completamente muerta, pero no es así. La inteligencia artificial (IA) en su versión fuerte es una expresión de la idea de que la mente no es parte del mundo natural. Uno de los adherentes de la visión fuerte de IA declaró que la mente es algo formal y abstracto (Dennett y Hofstadter 1981). La segunda tradición es la que erróneamente se designa como “materialismo”, sus seguidores frecuentemente niegan sin más que la conciencia sea cualitativa y que sus estados subjetivos existan realmente. Una declaración extrema fue dada por John B. Watson, uno de los fundadores del conductismo: “...ha llegado el tiempo de que la psicología descarte toda referencia a la conciencia... no es un concepto ni definible ni útil, es simplemente otro término para el ‘alma’ de los viejos tiempos...” (1925).

Ambas tradiciones comparten la resistencia a tratar a la conciencia real —la conciencia cualitativa y subjetiva— como un rasgo biológico real del mundo.

OBJETIVIDAD, SUBJETIVIDAD Y RELATIVIDAD DEL OBSERVADOR

Antes de refutar estas perspectivas y ofrecer una descripción de la conciencia, necesito hacer algunas distinciones. En primer lugar, la distinción entre objetividad y subjetividad se encuentra ampliamente arraigada en nuestra cultura intelectual, pero, por desgracia, es sistemáticamente ambigua porque fluctúa entre un sentido epistémico y un sentido ontológico (lo “epistémico” significa que tiene que ver con el conocimiento y lo “ontológico” significa que tiene que ver con la existencia).

Epistémicamente, la distinción radica en los tipos de afirmaciones, aquellas que pueden establecerse como un “hecho objetivo” y aquellas que no. Así, si asevero que Van Gogh murió en Francia, ello es un hecho epistémicamente objetivo que usted puede establecer como tal. Ahora, si yo asevero que Van Gogh fue mejor pintor que Gauguin, como la gente suele decir, ello es un asunto de “opinión subjetiva” que no puede establecerse como un hecho objetivo. Además de ésta, hay otra distinción ontológica subyacente respecto a los modos de existencia. Algunas entidades tienen un modo de existencia que es independiente de la actitud o la conciencia de cualquier sujeto; las montañas, las moléculas y las placas

tectónicas son ontológicamente objetivas en este sentido. Sin embargo, hay otras entidades cuya misma existencia depende de que sean experimentadas por un humano o un animal, en este sentido los dolores, las cosquillas y la comezón son ontológicamente subjetivos. ¿Por qué es importante esta distinción? Es de crucial importancia porque la subjetividad ontológica en cierto dominio no excluye la objetividad epistémica de la ciencia en el mismo dominio. Se puede tener una ciencia de la conciencia o de la vida mental en general que sea completamente adecuada y a la vez epistémicamente objetiva, incluso cuando todo el dominio sea ontológicamente subjetivo.

Relacionada con la distinción entre objetividad y subjetividad está aquella entre los rasgos del mundo cuya existencia depende de actitudes humanas, y los que existen independientemente de las actitudes. Denomino a la primera como "relativa al observador" y a la segunda como "independiente del observador" o "absoluta". Los fenómenos relativos al observador incluyen el dinero, los bienes, el matrimonio, los Estados nacionales, las universidades y las vacaciones de verano. Los fenómenos independientes del observador incluyen a las montañas, las moléculas, las galaxias y las placas tectónicas. En general, las ciencias naturales tratan con fenómenos independientes del observador y las ciencias sociales tratan con fenómenos relativos al observador. La relatividad del observador respecto al fenómeno introduce un elemento de subjetividad en la existencia misma del fenómeno. De tal modo que la existencia del dinero y del lenguaje, por ejemplo, es relativa al observador y por lo tanto contiene un elemento de subjetividad ontológica. Sin embargo, y esto es el punto crucial, como hemos visto, la subjetividad ontológica de un dominio no se opone a una ciencia epistémicamente objetiva del mismo dominio. Puede haber una ciencia epistémicamente objetiva de la economía y la lingüística aunque los objetos de estudio de esas disciplinas sean mayormente relativos al observador y así, al menos en parte, ontológicamente subjetivos. Todos los fenómenos relativos al observador son creados por la conciencia, pero la conciencia que los crea es absoluta y no es relativa al observador. Posteriormente diré algo respecto al papel del inconsciente.

Estas distinciones nos permiten ya responder a varias de las teorías erróneas acerca de la conciencia. Sin embargo, antes de hacerlo quiero abordar directamente el análisis de la conciencia.

ANÁLISIS DE LA CONCIENCIA

¿Qué es la conciencia? ¿Cuáles son sus características? ¿Dónde nos encontramos en su investigación científica? Suele decirse que la conciencia es difícil de definir, pero yo no lo creo. Necesitamos diferenciar entre la definición que resulta de la investigación científica y la definición de

sentido común que nos permite identificar el objeto de la investigación. Piénsese en el agua. La definición de sentido común es que es un líquido claro, incoloro e insípido que fluye en ríos, lagos y cae del cielo en forma de lluvia. La definición científica que resulta de la investigación es que: es H₂O. Con la conciencia estamos todavía en la etapa de líquido claro, incoloro, insípido y este es el tipo de definición que ofreceré. La conciencia consiste en todos nuestros estados de sensación, autoconciencia perceptiva, o estado de alerta. Comienza cuando despertamos de una ensoñación y continúa hasta que vamos a dormir nuevamente, o morimos, o, por el contrario, quedamos inconscientes. En esta definición, los sueños son una forma de conciencia.

Esta es entonces la definición de la conciencia. ¿Cuáles son algunas de sus características? Hay muchas características, pero para abreviar me limitaré a las cinco más importantes.

Cualitatividad. Para cada estado de conciencia hay una sensación cualitativa asociada, algo que se siente al estar en ese estado. Piense en la diferencia que hay entre beber vino, escuchar música y calcular sus impuestos. La cualitatividad no es sólo una característica de sensaciones corporales y emociones, sino que también caracteriza al pensamiento. Usted puede darse cuenta de que pensar que $2+2=4$ tiene cierta sensación característica; ahora trate de tener el mismo pensamiento en francés o alemán y, por lo menos para mí, se siente diferente.

Subjetividad ontológica. Debido a este carácter cualitativo, cada experiencia consciente es ontológicamente subjetiva, sólo existe en la medida en que es experimentada por un sujeto humano o animal. Esta característica hace que la conciencia parezca intratable; la hace parecer como un tema difícil para el estudio científico.

Unidad. Una tercera característica de la conciencia es que hay una notable unidad en nuestros estados de conciencia. Al dar una conferencia, no sólo escucho el sonido de mi voz, veo a la gente de la audiencia y siento un leve dolor de cabeza por el vino que bebí la noche anterior, sino que tengo todo esto como parte de un campo consciente unificado. Cuando lleguemos al tema de la evolución, veremos la unidad como una cuestión central que permite que la conciencia lleve a cabo sus funciones biológicas. Uno nunca tiene una experiencia consciente aislada como el sabor del vino o el olor de la rosa, sino que cada experiencia es parte de un amplio fenómeno consciente, que consiste en el campo ontológico subjetivo total de la conciencia en ese momento.

Solía pensar que la cualitatividad, la subjetividad y la unidad eran características separadas, pero ahora creo que estas son sólo diferentes aspectos de la misma característica, y esta es la esencia de la conciencia. No puedes tener cualitatividad sin subjetividad y no puedes tener subje-

tividad sin unidad. Esta característica es, por cierto, la razón por la cual los experimentos sobre escisión cerebral nos resultan tan interesantes: porque parece que la unidad de la conciencia se interrumpe parcialmente cuando se corta el cuerpo calloso. Estos son tres aspectos de una misma característica, y esta es la esencia de la conciencia. Hay otras dos características que quiero mencionar.

Intencionalidad. La intencionalidad es esa característica de la mente que se dirige a objetos y estados de cosas (la intención en su sentido ordinario es sólo un tipo de intencionalidad, junto con la percepción, la creencia, el deseo, el miedo, las emociones, y demás). No todos los estados conscientes son intencionales, piense por ejemplo en la ansiedad no dirigida. Sin embargo, la mayoría de ellos lo son, y esta característica es la que realmente hace importantes a las anteriores, porque la intencionalidad es la forma en que el organismo consciente se relaciona con el entorno.

Causalidad intencional. La mente consciente funciona causalmente porque produce el comportamiento y, en la percepción consciente, el entorno se presenta causalmente a la conciencia del agente. Por ejemplo, en presencia de alimentos, la luz que se refleja fuera del objeto causará una experiencia visual consciente en el animal, y esta experiencia motivará que el animal se mueva conscientemente hacia la comida.

SOLUCIÓN AL "PROBLEMA MENTE-CUERPO"

Supongamos que estoy en lo correcto hasta ahora y que la conciencia consiste en la subjetividad cualitativa unificada, y que esto nos permite lidiar con el entorno gracias a la intencionalidad y a la causalidad intencional. ¿Cómo encaja tal fenómeno con lo que conocemos del resto del universo? Usted no dudará en reconocer que este es el tradicional problema mente-cuerpo, que se supone que es terriblemente difícil, quizás imposible de resolver. Pienso que a nivel de las relaciones generales entre la mente y el cerebro, éste tiene una solución bastante simple. La solución neurobiológica resulta ser muy difícil y complicada, pero a nivel descriptivo, de las relaciones generales entre la conciencia y la neurobiología, la solución no es complicada. Y es esta: todos nuestros estados conscientes sin excepción son causados por procesos neurobiológicos cerebrales y ellos se llevan a cabo en el cerebro como rasgos de nivel superior. La conciencia es un rasgo del cerebro que es similar, por ejemplo, a la liquidez del agua característica del sistema de las moléculas de H₂O. No es algo que resulte evidente por las moléculas de H₂O; es la condición del sistema de moléculas que las contiene. De la misma manera, mi estado consciente es la condición que mi cerebro contiene. Y así como un cuerpo acuoso puede ir de un estado líquido a un estado sólido, dependiendo del comportamiento

de los microelementos —las moléculas de H₂O— así también el cerebro puede ir de un estado consciente a uno inconsciente dependiendo del comportamiento de los microelementos —las neuronas.

Si así es como funciona en general, entonces, ¿por qué no nos ponemos a trabajar en descubrir exactamente cómo funciona? ¿Por qué está progresando tan lentamente? No lo sé. Me pidieron escribir un artículo para la *Annual Review of Neuroscience* (Searle 2000) en el cual revisé gran parte de la literatura investigando la causalidad y el cómo se materializa la conciencia en el cerebro. Se me ocurrió que parte de la dificultad es que las técnicas de investigación son actualmente insuficientes para abordar los problemas que realmente nos preocupan. Éstas, usualmente, son técnicas de imagen —la MRI funcional es la más influyente, o algunas otras formas de escaneo también importantes— o registros de células individuales. La mayoría de los trabajos más interesantes, por ejemplo, de la competencia binocular, tienden a enfocarse en estructuras neurobiológicas muy locales, y a mí me parece que es más probable que la conciencia sea un fenómeno global. Nuestras técnicas de investigación actuales hacen extremadamente difícil investigar cómo la conciencia, como un fenómeno global, se produce en el cerebro.

Distingo entre lo que llamo el enfoque de “bloque fundamental”, donde concebimos al campo total consciente como constituido a partir de un conjunto de bloques conscientes separados, como los tipos de percepción específicos; y lo que llamo el enfoque de “campo unificado”, donde pensamos en estímulos específicos, como la percepción, como modificadores del campo consciente preexistente. Hubo un tiempo en el que parecía que la tarea crucial de la investigación era descubrir el correlato neuronal de la conciencia. Y de hecho en este periodo fueron descubiertos muchos correlatos neuronales de formas específicas de la conciencia. Sin embargo, estos descubrimientos no resolvieron el problema de la conciencia en la forma que algunas personas esperaban. ¿Por qué no? La investigación sobre la visión ciega, la competencia binocular, o los cambios gestálticos, se hizo en sujetos que estaban conscientes. Aun así, no respondieron la pregunta de, para empezar, ¿cómo el cerebro del sujeto crea un campo consciente completo? Existen dos visiones rivales de esta investigación: una es que la percepción crea a la conciencia; la segunda, a la cual yo me adhiero, es que deberíamos pensar en la percepción como modificadora del campo consciente preexistente. Y la pregunta crucial es: ¿cómo el cerebro crea el campo consciente?

Piénselo de la siguiente manera. Imagine que despierta en un cuarto totalmente oscuro en un ambiente desconocido; usted puede estar totalmente despierto y alerta, aunque tenga cero o mínimos estímulos perceptuales, si usted se levanta y se mueve, ¿está creando conciencia? En un sentido sí, lo hace porque tiene estados conscientes que no tenía antes. Sin

embargo, creo que es mejor pensar en estas experiencias como modificadoras del campo consciente preexistente que surgió cuando usted está totalmente despierto. La mayor parte de la investigación que conozco se basa ya sea en técnicas de imagen o en registros de actividad celular individual. La dificultad está en que ninguna de ellas parece adecuada para captar las propiedades globales del cerebro, y podría ser el caso de que la creación del campo consciente requiera de técnicas más poderosas que éstas.

REFUTACIÓN DE LAS PERSPECTIVAS ERRÓNEAS

Prometí hacer al menos una breve refutación de las seis teorías erróneas de la conciencia y ahora tenemos suficiente material para ello.

Conductismo. Debería avergonzarnos que el conductismo haya sido tan influyente por tanto tiempo porque obviamente es falso. Éste niega la subjetividad de la conciencia. Cada uno de nosotros sabe por experiencia propia que el dolor es una cosa y que el dolor observado es otra cosa.

Computacionalismo. Sabemos que un programa de computadora implementado no es por sí mismo suficiente para los procesos mentales, ya sean conscientes o inconscientes, porque el programa se define únicamente en términos sintácticos de manipulación del símbolos, mientras que los procesos mentales tienen contenido real. La sintaxis en sí misma no es suficiente para explicar el contenido semántico. Demostré esto hace una generación con el experimento mental (Searle 1980) de la habitación China. Imagine que está usted siguiendo los pasos de un programa para responder preguntas en chino o en algún otro idioma que usted no entiende; usted podría ofrecer respuestas correctas pero, de todas maneras, no entender chino. Seguir pasos computacionales no es suficiente para comprender.

Creo que el argumento es concluyente, pero el material de este artículo da para una discusión mucho mayor. A excepción de las computaciones llevadas a cabo por un agente consciente, el cómputo es relativo al observador. Usted no puede explicar la conciencia como una computación porque un proceso es computacional sólo con relación a algún agente consciente. Ya sea que un agente consciente esté llevando a cabo un cálculo, como la suma de $2+2$ para obtener 4, o que un agente consciente esté utilizando o pueda utilizar una máquina como una calculadora para poder interpretar los resultados como aritméticos. Tales cálculos son siempre relativos al observador. Téngase presente además que la relatividad del observador no implica la subjetividad epistémica. Es un hecho epistémicamente objetivo que estoy escribiendo esto usando el programa *Word*, y que el "pro-

grama *Word*" se implementa electrónicamente, pero no designa un fenómeno eléctrico.

Epifenomenalismo. Tenemos literalmente miles de años de experimentar la conciencia humana y animal como causa del comportamiento. El problema es explicar cómo puede hacerlo, dada su ontología subjetiva. Así es como funciona: considérese un acto simple como subir un brazo. Mi intención en acción hace que mi brazo suba. Sabemos que no importa lo que cause que mi brazo suba, debe provocarse la secreción de acetilcolina en las terminales del axón de las neuronas motoras. Sin acetilcolina, no subiría mi brazo. Esto significa que la intención de la acción consciente tiene que ser un fenómeno bioquímico. No hay manera de producir la secreción de acetilcolina a menos que se lleve a cabo en una estructura biológica. En el mismo caso, mi intención de la acción consciente tiene un nivel de descripción que es cualitativo, subjetivo, y es parte del campo consciente unificado que también es subjetivo, y hay otro nivel de descripción que consiste en un proceso neurobiológico cerebral.

Parte de nuestra dificultad para comprender este punto es que estamos atrapados en el vocabulario tradicional que contiene los errores tradicionales: el vocabulario de la mente y el cuerpo, y el dualismo y el materialismo. Lo que estoy intentando transmitir con este ejemplo tan simple es que, incluso para actividades conscientes simples como levantar un brazo, las categorías tradicionales son obsoletas porque debemos tener el concepto de evento único que incluye rasgos subjetivos, cualitativos, mentales y bioquímicos. Y este fenómeno tiene un carácter familiar, porque usted tiene el mismo fenómeno, el mismo sistema, con diferentes niveles de descripción. El motor de mi auto tiene un nivel de descripción, según el cual ocurren explosiones en el cilindro que impulsa el pistón; y otro nivel de descripción donde se oxidan las moléculas individuales de los hidrocarburos. Nos resulta difícil apreciar estos niveles a los que la mente se refiere porque uno de esos niveles de descripción tiene una historia sórdida. La tradición dualista ha dado mala fama al nivel mental de descripción porque hace parecer que nuestra vida mental no es parte de nuestra existencia biológica ordinaria.

Potencial para la acción. En estos experimentos se pidió a los sujetos que llevaran a cabo un acto trivial como apretar un botón y observar un reloj exactamente cuando comenzaran a hacerlo. Unos 200-300 ms antes de que se percataran de que habían decidido hacerlo, hubo un aumento de actividad en el área motora suplementaria. Se hicieron afirmaciones increíbles sobre estos datos, como por ejemplo, que refutaban el libre albedrío y demostraban que nuestro cerebro decide llevar a cabo acciones antes que nuestra mente consciente lo haga (Koch 2012). Los experimentos recientes demuestran que los experimentos originales eran defectuosos.

Si se le pide a los sujetos mirar al reloj y deciden no realizar tal acción, se obtiene el mismo potencial para la acción. Lo que podemos decir es que el potencial para la acción se produjo por mirar el reloj; si quitamos el reloj, no hay potencial para la acción (Trevena y Miller 2010).

Creo que la historia del potencial para la acción es un capítulo desafortunado en la historia científica reciente y plantea la pregunta: ¿por qué las personas estaban tan ávidas por creer tales conclusiones inverosímiles? La respuesta, creo, es que querían desacreditar a la conciencia. La conciencia en general ha sido una vergüenza para las ciencias naturales y, en este caso, parece que tendríamos pruebas científicas de que la conciencia realmente importa poco para nuestra conducta.

Argumento objetivo/subjetivo. Este argumento es obviamente una falacia por la ambigüedad en los sentidos objetivo y subjetivo que he explicado. La ciencia es, en realidad, epistémicamente objetiva. No obstante, no hay nada en la objetividad epistémica que impida la investigación de un dominio que es ontológicamente subjetivo.

Materialismo. Estas categorías de materialismo, mentalismo, dualismo, etc., se han vuelto obsoletas. Por supuesto, la realidad última es descrita por las ciencias naturales y por lo tanto es "material". No hay nada en este concepto material que impida que la conciencia subjetiva, cualitativa, sea un fenómeno biológico como la digestión, la mitosis o la fotosíntesis.

LA REVOLUCIÓN DARWINIANA

Una cosa notable respecto al desarrollo del conocimiento es que no sólo tenemos nuevas explicaciones, sino nuevas formas de explicación. Para mí, una de las cosas más fascinantes de la revolución darwiniana es que tenemos una forma de explicación que era desconocida o ciertamente poco apreciada. La idea era que, además del nivel de explicación de la biología aristotélica tradicional, donde teníamos una explicación teleológica del fenotipo, se substituyó por una explicación en dos niveles diferentes. Aristóteles pensaba que había tales cosas como causas finales o teleológicas, donde la explicación se daba especificando la meta, objetivo o fin del fenómeno a explicar. Así, si quisiera explicar por qué los peces tienen la forma que tienen, por qué los peces no tienen la forma de un ladrillo, o por qué las plantas inclinan sus hojas hacia el sol, se debería apuntar que el propósito de todo esto es que los peces naden mejor o que la planta sobreviva. Es tal meta teleológica la que provee la explicación. La revolución darwiniana produjo una substitución de dos diferentes niveles de explicación. En lugar de decir que la planta inclina sus hojas hacia el sol porque tiene la meta de sobrevivir, sustituimos dos niveles; una explicación causal o mecánica y una explicación funcional. En el nivel mecánico,

la planta tiene secreciones variables de la hormona de crecimiento auxina, y estas secreciones variables de auxina hacen inclinar las hojas hacia el sol. En el segundo nivel, el funcional, las plantas que inclinan sus hojas hacia el sol tienen más probabilidades de sobrevivir que las plantas, que no lo hacen. Nótese que la supervivencia todavía opera en la explicación, pero la supervivencia ya no es la meta que tiene la planta, es sólo algo que sucede. Así, hemos invertido el condicional. En lugar de decir que para sobrevivir la planta tiene que inclinar sus hojas hacia el sol, ahora decimos que la planta inclinará sus hojas hacia el sol debido a secreciones químicas; y puesto que inclina sus hojas hacia el sol, es más probable que sobreviva si lo hace que si no lo hiciera. Se sustituye la tradicional causa final aristotélica. La supervivencia todavía opera, pero ya no como la meta que explica al fenómeno, sino como algo que ocurre. Este rasgo introduce otro elemento a la explicación: el elemento diacrónico. Este tipo de explicación opera sólo con periodos de tiempo.

EL FRACASO DE LA SOCIOBIOLOGÍA

Hay límites respecto a lo que se puede hacer con los modos de explicación darwiniana. Hace algunas décadas, hubo un movimiento llamado sociobiología, el cual había adoptado este modelo de explicación para dar cuenta de la cultura humana, la moral, la civilización, la filosofía y casi todo lo demás. El inventor de esto fue E. O. Wilson (1975). Tuve la oportunidad de debatir con él y algunos otros autores, a veces por escrito y más frecuentemente en conferencias. El fracaso de la sociobiología es revelador. He defendido el modo de explicación darwiniano, pero, ¿cuáles son sus límites?, ¿por qué la sociobiología fracasó? El asumir que falló puede sonar a una petición de principio, por ello quiero explicar un poco más cuáles fueron sus limitaciones.

La limitación más cruda que tenían los métodos sociobiológicos era que trataban de explicar las características específicas de la sociedad y la cultura humana, de allí el nombre de sociobiología, pero el modo de explicación debía ser consistente con el hecho de que no ha habido ningún cambio importante en el genoma humano en los últimos 30,000 años. Esta cifra, por cierto, proviene de antropólogos físicos de la Universidad de California, Berkeley, especialmente Sherwood Washburn (tal vez las cifras son poco precisas, lo cual discutiré en un momento), pero, si vamos a explicar las sociedades humanas, piense en las enormes variaciones en esta sociedad en los últimos 30,000 años. De modo que si queremos explicar cosas como el surgimiento del fascismo, o la sociedad democrática que se produjo a partir de la Ilustración, o la existencia misma de la Ilustración, entonces parece que tenemos una herramienta analítica demasiado primitiva para trabajar. Supongamos que la cifra está mal, que está equivocada

por un factor de diez, y supongamos que el genoma humano es lo que es desde hace 3,000 años, aún así, el mismo problema se nos presenta. Hay todavía demasiada variación en los últimos 3,000 años como para que un único modo de explicación dé cuenta de todo esto.

Puede ser que la respuesta se concentre en los universales culturales, y eso es lo que hizo Wilson. Por ejemplo, a partir de que todas las sociedades tienen el tabú del incesto, Wilson pensó que la sociobiología proporcionaba una explicación para tal tabú. Los antropólogos culturales estaban algo indignados por esto, y señalaron, entre otras cosas, que el tabú del incesto toma muchas formas distintas en diferentes culturas. En algunas culturas está prohibido casarse con un primo y en otras culturas no lo está. Sin embargo, creo que la respuesta es que el tabú del incesto siempre prohíbe las relaciones sexuales entre hermano-hermana, padre-hija y madre-hijo, y éste es el núcleo del tabú. Una vez debatí con Wilson en la Universidad de Michigan; él afirmó que la sociobiología ha mostrado por qué el incesto es malo; por primera vez tenemos una explicación de una característica universal de la moralidad humana, tenemos una explicación de por qué el incesto es malo. ¿Cómo se supone que la explicación funcione en la analogía con la explicación que di de los dos niveles de explicación por los cuales las plantas inclinan sus hojas hacia el sol? Los estudios han mostrado que los niños criados cerca de personas del sexo opuesto carecen de deseo sexual hacia ellos. El ejemplo favorito viene de los *kibbutz* israelíes. Resulta que los niños criados en un *kibbutz* tienden a carecer de deseo sexual por otras personas de ese mismo *kibbutz*. Ellos se crían muy de cerca, en la estancia infantil. En general, tienden a tener mayores deseos sexuales por la gente que se crió en un *kibbutz* diferente, al otro lado de la colina. Wilson explicó que lo que pasa es que la cercanía durante el periodo de desarrollo genera un tipo de aversión; esta aversión sexual ofrece una explicación mecánica o causal del tabú del incesto. ¿Cuál es la explicación funcional? Para responder a esto, tenemos que preguntarnos, en primer lugar, ¿por qué tenemos reproducción biparental? La respuesta es que se obtiene un resultado genético mucho mejor si se mezclan los genes. Si sólo nos reprodujéramos como la amiba, por fisión o por incesto, no obtendríamos las ventajas de mezclar genes. Así, tenemos el nivel funcional de explicación: las ventajas de la reproducción biparental, muchas de las cuales se perderían si nos reprodujéramos por incesto. Y este es el nivel causal de la explicación, que es la aversión.

La anterior me parece una explicación inadecuada de una prohibición universal. Para empezar, la inhibición no explica la prohibición. Si fuera cierto que los seres humanos tienen una inhibición innata para relacionarse sexualmente con personas que han sido criadas de forma muy cercana, entonces habría poca necesidad para un tabú del incesto. No hay necesidad, por ejemplo, de una prohibición de comer lodo crudo, porque no

tenemos inclinación hacia ello. Sin embargo, la ausencia de una inclinación no explica el poder de la prohibición. Y el nivel funcional no da el contenido intencional específico para la prohibición. No vamos de “la mezcla de los genes es útil” a “el incesto es malo”. Además, cuando Wilson dijo que por fin habíamos explicado por qué el incesto es malo, tal explicación de hecho no formulada. Cuando mucho, podría ser una explicación de por qué es una mala idea tener relaciones incestuosas sin tomar las precauciones necesarias para prevenir la fecundación. No obstante, la idea de que el componente moral en el tabú del incesto se explica porque hay una ventaja funcional en la reproducción biparental y una inhibición natural que ocurre cuando la gente crece en proximidad, me parece demasiado débil. Este es un ejemplo de, finalmente, dónde fallan las formas sociobiológicas; fallan en describir la estructura general de la sociedad humana o la estructura de la moral. Esto no quiere decir que no podamos hacer más de lo que hemos hecho con los modos de explicación darwiniana, sino sólo que este movimiento intelectual en particular desde el principio me parecía que estaba condenado al fracaso.

LA FUNCIÓN EVOLUTIVA DE LA CONCIENCIA

Entonces, tal punto plantea la pregunta que a menudo escucho cuando hablo acerca de la conciencia: ¿cuál es la función evolutiva de la conciencia? Tal vez no tiene una función. Hay dos clases de personas con las que he debatido al respecto. Unas dicen que la conciencia debe ser epifenoménica, porque el mundo físico está causalmente cerrado, y en su descripción la conciencia es irreductiblemente subjetiva, y por lo tanto no puede reducirse al mundo objetivo. Ya he dado respuesta a esta afirmación. La segunda forma de argumentación es más desafiante. La conciencia no tiene una función porque podemos fácilmente imaginar que toda la vida humana y animal seguirán siendo como hasta hoy, pero ahora, sin la conciencia. Podemos fácilmente imaginar que somos como zombies inconscientes que llevábamos a cabo exactamente las mismas acciones que llevamos a cabo ahora. Esta posibilidad mostraría que la conciencia no tiene en esencia ninguna función evolutiva (con frecuencia escuchamos esta objeción en conferencias sobre la conciencia; recientemente me encontré con esta objeción hecha por varios críticos en el ciclo de Conferencias Turing 2012 sobre la Conciencia en la Universidad de Québec en Montréal). Este es un error mucho más profundo que responderé.

Solemos escuchar la pregunta, “¿cuál es la función evolutiva de la conciencia?, ¿qué ventaja selectiva transmite?”, en analogía a preguntas como: “¿cuál es la ventaja selectiva del reflejo ocular vestibular (ROV)?, ¿qué ventaja selectiva conlleva?” En el caso del ROV, la respuesta es simple: al estabilizar la imagen retinal, mejora la percepción visual. Llegamos a

esta respuesta sustrayendo mentalmente el ROV y viendo qué diferencia haría. El problema de tratar de hacer esto con la conciencia es que, si se sustrae a la conciencia, se está sustrayendo, en términos generales, a toda nuestra vida, excepto por algunos procesos básicos e inconscientes como respirar. ¿Qué se perdería si de repente entráramos en coma o siempre hubiésemos estado en coma? En términos generales, todo. La razón por la que muchas personas no ven esto es porque normalmente escuchan la pregunta, “¿qué se perderían si se sustrajese la conciencia?” o “¿qué perderíamos si se sustrajese la conciencia, pero dejamos todo lo demás igual”? No obstante, por supuesto, todo lo demás no podría ser lo mismo. De hecho, prácticamente nada sería lo mismo. Yo no podría levantarme de la cama por la mañana, ni comer, ni rondar por ahí en la Tierra, ni tratar con otras personas. Incluso es un error pensar que la conciencia tiene “una función”. Es como si se preguntara: ¿cuál es la función de la vida? Todo en nuestra vida depende de estar vivos y casi todo en nuestras vidas depende de estar conscientes. La razón por la cual la gente tiende a pensar que podemos sustraer mentalmente a la conciencia y dejar todo lo demás igual es porque son víctimas de dualismo residual. Asumen que la conciencia no es parte del mundo “físico” ordinario, como si pudiéramos quitarla y dejar todo lo demás igual. Así no funciona la naturaleza. De modo que esto no muestra que las alas de los pájaros no tengan una función, o argumentar que somos capaces de imaginar un mundo en el que los pájaros son impulsados por el motor de un cohete. El hecho de que podamos imaginar a zombis que se comportan como nosotros pero sin la conciencia, es como imaginar que los pájaros son impulsados por el motor de un cohete. Esto es irrelevante para una explicación evolucionista. La manera en que funciona la naturaleza para seres como nosotros es que tenemos un enorme poder añadido gracias a la existencia de la conciencia, a la existencia de estados conscientes que son subjetivos y cualitativos.

Nuestro vocabulario común hace difícil establecer o apreciar este punto porque tendemos a pensar en términos de “información” y “representación”, y es tentador decir cosas como “tenemos un procesamiento de información mucho más eficiente y formas de representación más efectivas que las que tendríamos si fuéramos inconscientes”. Esta formulación no hace justicia al poder de la conciencia. Nuestras percepciones y acciones conscientes no sólo nos dan representaciones, sino que nos dan representaciones directas del entorno. Gracias a la conciencia, me percato inmediatamente del ambiente a mí alrededor. Y gracias a la unificación del campo consciente, tengo una prodigiosa cantidad de contenido representacional en el campo consciente. Si usted mira a su alrededor por un momento y piensa en de cuántas cosas puede conscientemente percatarse en este momento, nunca sabría donde detenerse. Y recuerde que usted no sólo tiene el carácter de representación inmediata, sino que además tiene

la sensación inmediata de su pasado más reciente y de su posible futuro. Usted constantemente está utilizando la unificación del campo consciente para realizar intercambios con el entorno.

Todavía no he expresado adecuadamente el enorme poder que nos da la conciencia. El organismo es capaz de coordinar un enorme número de estímulos simultáneos dentro de un único campo consciente y también es capaz de coordinar su comportamiento a la luz de los insumos sensoriales y a la luz de sus metas y los medios posibles para lograrlas. Coordina todo esto de manera que simultáneamente representa el pasado en forma de memoria de corto y de largo plazo, y anticipa el futuro a partir de sus planes, metas e intenciones. En el caso de los seres humanos, la conciencia también nos da la capacidad para cooperar con nuestros semejantes, y la forma específica que toma la cooperación humana es el uso consciente del lenguaje. El uso del lenguaje nos da los rasgos distintivos de la civilización humana: dinero, propiedad, gobierno, organizaciones sociales, etc., todos ellos resultados de la aplicación consciente de las representaciones lingüísticas.

EL INCONSCIENTE

¿Qué deberíamos decir sobre el inconsciente? Es esencial distinguir entre los fenómenos mentales inconscientes y los procesos neurobiológicos no conscientes que posibilitan la conciencia. Por ejemplo, incluso cuando estoy profundamente dormido tengo la creencia inconsciente de que George Washington fue el primer presidente. Cuando veo algo, hay mecanismos de retroalimentación no consciente entre el núcleo geniculado lateral y el V1. Estos son esenciales para la percepción, pero no tienen una realidad psicológica en absoluto. ¿Cuál es la diferencia? La razón por la cual podemos decir que mis creencias inconscientes son psicológicamente reales es que soy capaz de traerlas a la conciencia. De este modo, la noción de un estado mental inconsciente es parasitaria de la noción de conciencia porque una creencia inconsciente, que efectivamente tiene una realidad psicológica, tiene que ser el tipo de cosa que en principio se podría traer a la conciencia. Tengo que decir "en principio", porque el acceso a la conciencia podría estar bloqueado por un daño cerebral, o una represión, o el olvido, u otro tipo de cosas, pero no podría ser algo que no fuera en principio accesible a la conciencia porque, de ser así, no tendría ninguna realidad psicológica.

Es importante destacar este punto porque en gran medida, en los inicios de la ciencia cognitiva, el paradigma explicativo consistía en intentar tener teorías computacionales de las capacidades cognitivas (Marr, et al. 2010). Se nos dijo que el nivel computacional tenía realidad psicológica, pero está a un nivel intermedio entre la psicología del sentido común, consciente y

similares, y el nivel de la neurobiología. Afirmo que no existe tal nivel. No tiene sentido la idea de que hay una realidad psicológica a ese nivel si no posee las características que, en principio, forman parte de la accesibilidad a la conciencia.

Considero que el paradigma de la ciencia cognitiva está cambiando. Pienso que nos estamos alejando de una ciencia cognitiva computacional a una neurociencia cognitiva. Doy la bienvenida a ese movimiento. Pienso que el futuro de este proyecto de investigación completo está en una mejor comprensión del cerebro.

REFERENCIAS BIBLIOGRÁFICAS

- Chalmers, D. (1996), *The Conscious Mind: In Search of a Fundamental Theory*. New York: Oxford Univ. Press.
- Dennett, C. D. (1991), *Consciousness Explained*. Boston: Little, Brown.
- Dennett, C. D. y R. D. Hofstadter (1981), *The Mind's I: Fantasies and Reflections on the Self and Soul*. New York: Bantam House.
- Descartes, R. (1996), *Meditations on First Philosophy*. Trad. J. Cottingham. Cambridge, UK: Cambridge Univ Press. [Versión castellana, Descartes, R. (1987), *Meditaciones metafísicas*. Madrid: Gredos. Trad. López y M. Grana.]
- Koch, C. (2012), *Consciousness: Confessions of a Romantic Reductionist*. Cambridge, M. A.: MIT Press.
- Libet, B., C. A. Gleason, E. W. Wright y D. K. Pearl (1983), "Time of conscious intention to act in relation to onset of cerebral activity (readiness-potential): The unconscious initiation of a freely voluntary act," *Brain* 106 (Pt 3): 623–642.
- Marr, D., A. T. Poggio y S. Ullman (2010), *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. Cambridge, M. A.: MIT Press.
- Minsky, M. y S. Papert (1987), *Society of Mind*. New York: Simon and Schuster.
- Searle, J. R. (1980), "Minds, brain and programs," *Behav. Brain Sci.* 3(3): 417–457.
- Searle, J. R. (1992), *The Rediscovery of the Mind*. Cambridge, M. A.: MIT Press. [Versión castellana, Searle J. R. (1996), *El redescubrimiento de la mente*. Barcelona: Crítica.]
- Searle, J. R. (1996), *Minds, Brains and Science*. Cambridge, M. A.: Harvard Univ. Press. [Versión castellana, Searle, J. R. (1985), *Mentes, cerebros y ciencia*. Ediciones Cátedra.]
- Searle, J. R. (2000), "Consciousness," *Annu. Rev. Neurosci.* 23: 557–578.
- Searle, J. R. (2005), *Mind: A Brief Introduction*. Oxford: Oxford Univ. Press. [Versión castellana, Searle, J. R. (2006), *La mente: una breve introducción*. Grupo Editorial Norma: Bogotá. Trad. Horacio Pons]
- Skinner, B. F. (1992), *Verbal Behavior*. Acton, M. A.: Copley Publishing Group.
- Trevena, J. y J. Miller (2010), "Brain preparation before a voluntary action: Evidence against unconscious movement initiation," *Conscious Cogn.* 19(1): 447–456.
- Watson, J. B. (1925), *Behaviorism* (New York: Norton); citado en B. J. Baars (2003), "The double life of B. F. Skinner," *J. Conscious Stud.* 10(1): 5–25.
- Wilson, E. O. (1975), *Sociobiology: The New Synthesis*. Cambridge, M. A.: Belknap Press of Harvard Univ.