

Claudio Gutiérrez

## La explicación de la mente: A propósito de un libro de Daniel Dennett

**Summary:** *Daniel Dennett, in his new book *Consciousness Explained*, attempts to clarify important philosophy of mind issues by applying to them scientific and technological criteria.*

*On the first attempt, he notes that explaining a phenomenon means to get rid of the original puzzle: consciousness is explained on the basis of things which are not conscious. On the second attempt, he presents impressive evidence about how much one can clarify a philosophical problem by applying to it the engineering criterion of what is possible to build.*

*The final upshot of Dennett's analysis is the proposal of a new model for the mind, the Multiple Draft model, to replace the dualist traditional one: the Cartesian Theater model.*

**Resumen:** *Daniel Dennett, en su nueva obra *La explicación de la conciencia*, intenta aclarar importantes asuntos de filosofía de la mente mediante la aplicación de criterios científicos y tecnológicos.*

*En cuanto al primer intento, hace notar que explicar un fenómeno significa librarse del asombro original: la conciencia tiene que explicarse en términos de cosas que no son conscientes. En cuanto al segundo intento, presenta pruebas impresionantes de cómo es posible clarificar un problema filosófico por la aplicación del criterio ingenieril de lo que es posible construir.*

*El resultado final del análisis de Dennett es la propuesta de un nuevo modelo de la mente, el Modelo de las versiones múltiples, que reemplaza al dualista tradicional: el Modelo del teatro cartesiano.*

Un misterio es un fenómeno en relación con el cual la gente no ha aprendido a pensar todavía. Daniel Dennett

La nueva filosofía de la mente, en contraste con la tradicional, se caracteriza por la aplicación al problema de comprender los hechos mentales del punto de vista del método científico y el punto de vista del quehacer tecnológico. En otras palabras, se trata de producir explicaciones que despejen el "misterio" tradicional en los problemas de la mente y de acercarse a los fenómenos mentales con la misma actitud con que enfrentaríamos el funcionamiento de un artefacto cuyo diseño interior no conocemos.

### La mente puede ser enfocada con sentido científico

Sólo una teoría que dé cuenta de los eventos conscientes en términos de eventos no conscientes puede dar cuenta en absoluto de la conciencia, así como los sólidos y los líquidos y los gases no pueden ser explicados sino en términos de cosas que no son ellas mismas ni sólidos ni líquidos ni gases. Explicar el dolor sin despejar su carácter desagradable sería circular- tanto como explicar la acción del opio por su *virtud dormitiva* es circular, como ya lo comprendió Molière. Similarmente, una explicación adecuada de la risa debe dejar fuera la comicidad: su presencia meramente pospondría el intento de contestar la pregunta por su naturaleza. Si este enfoque nos parece extraño, como observa Daniel Dennett, ello sólo significa que no hemos aprendido aun a pensar sobre la mente, como sí sabemos ya pensar sobre los sólidos, los líquidos y los gases. En el fondo, nunca llegamos a *entender* una explicación

científica de un fenómeno natural: simplemente *nos acostumbramos* a ella, dadas las ventajas inmensas de carácter pragmático que representa el aceptarla. Si además de seres prácticos somos poetas, o místicos, siempre podemos poner temporalmente entre paréntesis la explicación científica y volver a maravillarnos de que el sol "salga" por el este todas las mañanas.

### La mente puede ser enfocada con sentido tecnológico

En cuanto al enfoque tecnológico, poner atención al fenómeno mental de la misma manera en que atenderíamos a un artefacto, permite el surgimiento de preguntas nuevas más fácilmente contestables que las de la filosofía tradicional, tales como: ¿cuáles modelos de la alucinación pueden evitar la explosión combinatoria implicada en suponer que el mecanismo alucinatorio debe prever las reacciones del alucinado? ¿cómo puede el contenido de la experiencia ser elaborado por procesos (relativamente) desprovistos de inteligencia y comprensión, del mismo modo que la música puede salir de un choque de partículas en el disco de un gramófono? En cuanto a la primera, podremos fácilmente mostrar que las alucinaciones fuertes son simplemente computacionalmente imposibles. Una alucinación fuerte sería, por ejemplo, las producidas por el "genio maligno" que Descartes supone metodológicamente en sus *Meditaciones metafísicas* para explicar las impresiones que recibimos del mundo exterior; o un fantasma con el que pudiéramos conversar, que nos permitiera tocarlo y resistiera nuestro embate con solidez, que proyectara una sombra como los cuerpos ordinarios, que fuera visible desde cualquier ángulo, que lo encontraríamos cómodamente sentado en el mismo sillón después de ausentarnos brevemente de la sala en que nos hubiere aparecido, etc. La respuesta tecnológica al primer interrogante es que sólo un fantasma verdadero podría producir una alucinación de ese tipo, y sólo un mundo real podría crear en la mente las impresiones inducidas por el genio maligno de Descartes. En cuanto a la segunda pregunta, sobre cómo producir cognición con elementos no cognitivos, puede considerarse la como el programa de investigación mismo de la moderna ciencia del conocimiento. Parte de ese programa, por ejemplo, implica explicar el senti-

do común de la gente por la interacción en el cerebro de sistemas "menos inteligentes" que colaboran para producir la inteligencia ordinaria. Esos sistemas menos inteligentes, quedarían explicados por la coalición de subsistemas todavía menos inteligentes, hasta llegar a *átomos de inteligencia* ninguna de cuyas partes tuviera ya inteligencia<sup>1</sup>.

### La debilidad de la intuición

Uno de los obstáculos para la explicación científica de la inteligencia estriba en la tendencia inveterada que tenemos de creer que, por ser nosotros conciencia, tenemos una vía privada y privilegiada al conocimiento de sus secretos. En otra parte he llamado a esa tendencia "la tentación del científico social", y he tratado de exorcizarla mediante un cuidadoso distingo entre dos conceptos de subjetividad: *subjetividad 1* y *subjetividad 2*, según nos estemos refiriendo a la subjetividad *del científico* o a la subjetividad *del objeto de ciencia*. En particular, cuando en relatos sobre experiencias subjetivas pretendemos estar usando no más que nuestros poderes de intuición u observación interna, en realidad nos encontramos embarcados en procesos de teorización improvisada y las más de las veces carentes de un mínimo control experimental. La verdad es que la gente se equivoca continuamente, aun dentro de contextos de experimentación controlada, en sus opiniones intuitivas sobre lo que hace o sobre la forma en que lo hace. No se trata de que mienta propiamente, sino de que *confabula*: llena lagunas, adivina, especula, y confunde la teoría con la observación. Hay amplia evidencia experimental, recogida en multitud de experimentos psicológicos, de que uno no tiene realmente autoridad para dar cuenta de lo que le esté pasando a uno mismo; solamente la tiene para dar cuenta de lo que *le parece* que le esté pasando.

### El método heterofenomenológico

Esta comprobación da base a Dennett para presentar su método heterofenomenológico como un método neutral para investigar y describir la fenomenología. Implica la extracción y purificación de textos de sujetos (aparentemente) parlantes, y el uso de esos textos para generar una fic-

ción teórica, el mundo heterofenomenológico del sujeto. Tal mundo ficticio queda poblado con todas las imágenes, eventos, sonidos, olores, sentimientos y sentimientos que el sujeto cree con sinceridad que existen. En su máxima extensión, es un retrato neutral de exactamente *lo que es ser ese sujeto* -en sus propios términos. Una vez extraída tal heterofenomenología, el teórico puede entonces pasar a la cuestión de lo que podría explicar la *existencia* de esta heterofenomenología en todos sus detalles. Los hechos sobre lo que la gente cree, e informa cuando expresa sus creencias, son fenómenos que una teoría de la mente debe explicar. La cuestión de si los ítemes así representados existen como objetos, eventos y estados reales en el cerebro es una cuestión empírica que debe ser investigada. Si se descubren candidatos reales adecuados, podemos identificarlos como los referentes de los términos del sujeto; de lo contrario, de todos modos tendremos que explicar por qué los sujetos creen que tales términos se refieren a algo existente.

### Observación e interioridad

Algunas de las paradojas de la conciencia que producen más perplejidad surgen cuando nos aferramos demasiado tiempo a buenos hábitos de pensamiento, hábitos que usualmente nos evitan meternos en problemas. Así, por ejemplo, siempre que lidiamos con una mente consciente, inmediatamente postulamos la existencia de un *punto de vista*, desde el cual tal mente se supone que percibe el mundo que la rodea. Una mente consciente es un observador, que acepta una porción limitada de toda la información disponible, desde un ángulo particular que determina cómo se hace esa selección de información. ¿Qué sucede, sin embargo, cuando nos concentramos en el observador mismo, y seguimos pensando que *dentro de él* hay también un punto de vista, es decir si tratamos de localizar el punto de vista del observador más precisamente como un punto interno a ese individuo? La presuposición simple que funciona tan bien en la escala mayor comienza ahora a decepcionarnos: tendería a indicarnos que existe un observador interno, pero la investigación empírica nos demuestra que no hay un punto único en el cerebro hacia donde toda la información se canalice. Esta comprobación tiene consecuencias muy vastas, que no son ni

lejanamente obvias y que pueden incluso resultar contraintuitivas. No obstante, son perfectamente compatibles con la marcha general del método científico que busca explicar los fenómenos que no comprendemos en términos de ideas ya conocidas: el fenómeno de la observación tendrá por fuerza que ser explicado por fenómenos que en sí mismos no son observaciones.

### El teatro cartesiano

Dennett llama a la idea de que existe un lugar central en el cerebro *materialismo cartesiano*: es la opinión a que uno llega cuando descarta el dualismo de Descartes pero omite descartar la imaginación de un teatro central (aunque material) donde "todo se arme". El materialismo cartesiano es la opinión de que existe una frontera en algún lugar del cerebro, la cual marca el orden de llegada en la "presentación" de la experiencia. El teatro cartesiano es un cuadro metafórico de cómo la experiencia consciente puede producirse en el cerebro. Parece al principio una extrapolación inocente del hecho familiar e innegable de que *para intervalos de tiempo cotidianos y macroscópicos* se puede en verdad ordenar los eventos en las dos categorías "lo no observado todavía" y "lo ya observado". Pero la idea de un centro especial en el cerebro es más bien una mala idea que confunde tenazmente nuestros intentos de pensar sobre la conciencia.

Una vez que descartamos la fantasmagórica *res cogitans* de Descartes, no hay ya ningún papel para un estación central o centro *funcional* en el cerebro. La glándula pineal no sólo no es ya el *fax del alma*, sino que tampoco es la oficina presidencial del cerebro; y por supuesto, tampoco lo es ninguna otra porción del cerebro. El cerebro es la Presidencia, pero no hay ninguna razón para creer que el cerebro mismo tenga ninguna otra presidencia más profunda, ningún santuario íntimo el arribar al cual constituya la experiencia consciente. En breve, *no hay ningún homúnculo dentro del hombre, ningún "cerebrúnculo" dentro del cerebro.*

### El modelo de las versiones múltiples

De acuerdo con el modelo de la conciencia que Dennett propone para sustituir al del teatro

cartesiano -y que llama *modelo de las versiones múltiples*- toda variedad de percepción, de pensamiento o actividad mental se realiza en el cerebro por medio de procesos de interpretación y elaboración paralelos y de múltiples vías. Para usar una metáfora útil, la información que entra en el sistema nervioso es sometida en el cerebro a una "revisión editorial" continuada, como la que se haría en las oficinas de un periódico.

La gente frecuentemente se sorprende al saber que en condiciones normales sus ojos avanzan en rápidas sacudidas de aproximadamente cinco fijaciones por segundo, y que este movimiento, como el movimiento de la cabeza, se elimina muy pronto por corrección editorial en el trámite informático desde el ojo hasta la conciencia. Los psicólogos han también descubierto algunos efectos especiales, como la interpretación de profundidad en estereogramas de puntos aleatorios, que confirman la existencia de correcciones editoriales muy tempranas. Estos procesos editoriales toman fracciones largas de segundo, durante las cuales pueden ocurrir varios agregados, incorporaciones, enmiendas y hasta sustituciones de contenido, en distintos órdenes temporales.

Contrariamente a lo que pretendieron los filósofos empiristas o positivistas, no experimentamos *directamente* nada. Lo que experimentamos es un producto de muchos procesos de interpretación- procesos editoriales, en efecto. Estos procesos, que tienen lugar por flujos de actividad en varias partes del cerebro, reciben representaciones relativamente crudas y sesgadas, y producen representaciones agregadas, revisadas y aumentadas. Como en el caso del trabajo editorial de un periódico, dónde termina la corrección "anterior a la publicación" y dónde comienza la "fe de erratas" es una distinción que solo puede trazarse arbitrariamente.

No existe un solo y definitivo "flujo de conciencia", no hay oficinas centrales, ni teatro cartesiano donde "todo se integre" para ser ojeado por un "significador" central. En vez de un flujo de conciencia único hay múltiples canales en los cuales los circuitos especialistas, en verdadero pandemonio, tratan de hacer al mismo tiempo sus varias cosas, creando sobre la marcha versiones múltiples de la información que podemos llamar *narraciones*. La mayor parte de estas narraciones fragmentarias desempeñan papeles de corta duración, pero algunas son promovidas a ulteriores papeles funcionales por la actividad de lo que Dennett denomina la *máquina virtual del cerebro*.

## La fijación del contenido de la conciencia

Como hace notar muy sagazmente Dennett, la detección o discriminación de aspectos o detalles no deben ser hechas por el cerebro *más que una sola vez*. Esto es, una vez que una "observación" particular de algún detalle o aspecto ha sido realizada, por una porción especializada y localizada del cerebro, el contenido de información así fijado no tiene que ser enviado a ninguna parte para ser rediscriminado por un discriminador "jefe". En otras palabras, la discriminación no lleva a una *representación* del elemento discriminado, para beneficio de la audiencia del teatro cartesiano pues simplemente no hay teatro cartesiano. Estas fijaciones de contenido, espacial y temporalmente distribuidas en el cerebro, son localizables con precisión tanto en el espacio como en el tiempo. Pero es siempre una cuestión abierta si un contenido particular así discriminado aparecerá eventualmente como un elemento de la experiencia consciente, y solo por confusión podemos preguntar *cuándo* llega a ser consciente.

En todo momento hay "versiones" múltiples de fragmentos narrativos en varios estadios de edición en distintos lugares del cerebro. El modelo de las versiones múltiples evita la tentación de suponer que debe haber una narración única (la versión "final" o "publicada") que sea canónica. Sondeos del flujo de conciencia a distintos intervalos, pueden producir diferentes efectos, al precipitar diferentes narraciones: si uno demora el sondeo demasiado, el resultado puede no ser ya narración del todo por haber sido eliminado por el proceso editorial el elemento narrado; pero si uno sondea "demasiado temprano", puede recoger datos sobre cuán tempranamente se logra una discriminación particular en el flujo de conciencia, aunque al costo de perturbar el progreso normal de ese flujo.

## Narración y sondeo

Aquello de lo que somos conscientes dentro de una cierta duración de tiempo no se define independientemente de los sondeos usados para precipitar narraciones sobre ese período. Todos hemos experimentado el fenómeno de conducir un vehículo por varios kilómetros mientras estamos sumergidos en conversación (o silencioso soliloquio), y de repente descubrir que no tene-

mos del todo recuerdo de la carretera, del tráfico, de las propias actividades de manejo. ¿Estábamos realmente inconscientes de todos los carros que pasaron, de los semáforos, de las curvas de la carretera? Con toda seguridad que no: si *hubiéramos sido sondeados*, en varios momentos del camino, sobre lo que habríamos visto *justo antes*, habríamos tenido por supuesto algo que contar.

## Refutación del teatro cartesiano

Una de las más poderosas pruebas de la inexistencia del teatro cartesiano puede ser discernida en la base psicológica del cine: el llamado fenómeno *phi*. En su caso más simple, el fenómeno consiste en que si dos o más manchas separadas por, a lo más, 4 grados de ángulo visual se encienden brevemente en sucesión rápida, se ve una sola mancha que parece moverse reiteradamente en viaje de ida y vuelta.

En un experimento de Paul Kolers y Michael von Grünau, estimulado por Nelson Goodman, se encendieron dos manchas de distinto color durante 150 milisegundos cada una (dentro de un intervalo de 50 milisegundos); la primera mancha pareció comenzar a moverse y *entonces* cambiar de color abruptamente a la *mitad de su ilusorio camino* hacia la segunda localización. Goodman se preguntó: "¿cómo es que el cerebro puede llenar la mancha en los momentos-lugares intermediarios a lo largo del sendero del primero al segundo centelleo *antes de que el segundo centelleo ocurra?*".

El modelo de las versiones múltiples está de acuerdo con Goodman en que el cerebro crea retrospectivamente el contenido (el juicio) de que hay movimiento intermedio, y en que este contenido deja su marca en la memoria. Pero el modelo afirma además que el cerebro no necesita "construir" ninguna representación para "rellenar" los blancos. Esto sería una pérdida de tiempo. El juicio está *ya formado*, así que el cerebro puede dedicarse a otras tareas. Una vez que una discriminación ha sido hecha, no tiene que volver a hacerse: el cerebro simplemente se ajusta a la conclusión lograda y usa la nueva interpretación de la información para modular la conducta subsiguiente.

Cuando se descarta el dualismo cartesiano, deben descartarse también el espectáculo que tendría lugar en el teatro cartesiano y su público,

pues en el cerebro no se encuentran ni espectáculo ni público algunos, y sólo allí deberíamos buscarlos.

## Un experimento revelador

Cuando muchas cosas suceden en muy poco tiempo, el cerebro debe hacer suposiciones simplificadoras para dar cuenta de los hechos. Veamos un ejemplo. Se aplican en ritmo unos golpecitos con martilletos mecánicos puestos en dos o tres localizaciones a lo largo del brazo, hasta con un pie de separación; por ejemplo, cinco en la muñeca seguidos por dos cerca del codo y después tres más en el brazo superior. Los golpecitos son aplicados con intervalos entre estímulos de entre 50 y 200 milisegundos. Así, un tren de golpecitos podría durar menos de un segundo, o tanto como dos o tres. El sorprendente efecto es que los golpecitos parecen a los sujetos viajar en secuencia regular sobre puntos equidistantes hacia arriba como si un animalito fuera saltando a lo largo del brazo.

Ante resultados experimentales como estos, uno tiende a preguntarse: *¿cómo sabe el cerebro* que después de cinco golpecitos en la muñeca va a haber unos golpecitos cerca del codo? Los sujetos experimentan la "salida" de los golpecitos de la muñeca comenzando con el segundo, y sin embargo en pruebas de control en las que los golpecitos del codo nunca llegan a producirse, los sujetos sienten todos los cinco golpecitos de la muñeca *en la misma muñeca*.

El cambio en el espacio (a lo largo del brazo) lo discrimina el cerebro sobre el tiempo. El número de golpecitos se discrimina también. Aunque en la realidad física los golpecitos están agrupados en localizaciones particulares, la suposición simplificadora es que se distribuyen regularmente sobre la extensión espacio-temporal de la experiencia. El cerebro descansa en esta interpretación parsimoniosa pero equivocada *después* que los golpecitos se registran, por supuesto, y esto tiene el efecto de editar (borrar) interpretaciones (parciales) anteriores.

Esto no debiera parecer tan extraño: no representamos los objetos vistos como existentes en la retina, sino más bien a distintas distancias en el mundo externo; lo mismo hacemos con los objetos percibidos por los oídos. ¿Por qué no podría el cerebro representar también los eventos

en el *punto temporal* donde tenga mayor sentido ambiental que sucedan?

### El momento de la conciencia

Siendo así que la distancia del dedo del pie al cerebro es mucho más grande que la de la frente al cerebro, los estímulos aplicados simultáneamente a estos sitios diferentes arribarán a la Presidencia en sucesión descompasada, si suponemos constante la velocidad de viaje a lo largo de todas las sendas. ¿Cómo hace el cerebro, podría alguien preguntarse, para "asegurar simultaneidad central"? Pero es un error suponer que el cerebro necesite resolver este problema, por una razón obvia de ingeniería: sería malgastar tiempo precioso conceder al rango completo de las operaciones un itinerario de "peor caso". ¿Por qué deberían señales vitalmente importantes procedentes de la frente demorarse *siempre* en una sala de espera solo porque *algún día* podría haber ocasión de que señales concurrentes de los dedos de los pies convergieran con ellas? En cambio, lo que sí tiene que asegurar el cerebro es que, en las circunstancias requeridas, la narración obtenga "sentido ambiental", de modo que, por ejemplo, un efecto no preceda a su causa. En otras palabras, no es tan importante la cronología de las señales atinentes como su organización, original o "editada", y lo único necesario es que esa organización dé base adecuada para el control del comportamiento.

Puesto que la cognición y el control -y por lo tanto la conciencia- están distribuidos a lo largo y ancho del cerebro, se sigue que ningún tiempo puede ser señalado como el preciso momento cuándo un evento se hace consciente. Como los sondeos precipitan narraciones, el momento mismo en que se realizan pueden tener efectos revisionistas de gran envergadura en los sistemas representativos usados por el cerebro.

En conclusión, y con base en el análisis de Dennett, podemos formular el *principio psicológico de incertidumbre*:

No podemos atribuir a un evento dado un momento específico en el flujo de conciencia independientemente de los sondeos particulares que hacemos para determinarlo.

### La superimposición de la conciencia

Las funciones múltiples para un mismo órgano, relativamente raras entre los artefactos, abundan en cambio en la Madre Naturaleza. Una de las razones por las cuales los teóricos de la mente tienen tanta dificultad para discernir en el cerebro un diseño plausible para la conciencia es el hecho de que han tendido a pensar que los elementos cerebrales sirven *cada uno una y sólo una función*. El resultado es que está siendo tremendamente difícil escudriñar cómo funciona el cerebro, precisamente porque fue diseñado por un proceso capaz de utilizar funcionalidad múltiple y superimpuesta, algo sistemáticamente difícil de discernir por la ingeniería inversa.<sup>2</sup>

La lentitud de la actividad mental consciente hace evidente que el cerebro no fue realmente diseñado -alambrado- para tal actividad. Los especialistas básicos de la conciencia son parte de nuestra herencia animal: no fueron diseñados para realizar acciones particularmente humanas, como leer o escribir, sino para agacharse en caso de peligro, evitar depredadores, reconocer caras, agarrar, lanzar, recoger bayas y otras tareas esenciales. Los enlistó la evolución, de modo oportunista, para nuevos papeles, que solo podían corresponder por casualidad y aproximadamente a sus talentos naturales.

Así por ejemplo, nuestros sistemas visuales, como los de muchos otros animales superiores, incluso los peces, son exquisitamente sensibles a patrones con un eje vertical de simetría. Esto se debe probablemente a que en el mundo natural de nuestros remotos antepasados las únicas cosas en la naturaleza que mostraban ejes de simetría vertical eran otros animales, y solamente estando de frente. Así, nuestros antepasados llegaron a equiparse con un sistema de alarma muy valioso, el cual se disparaba (sobre todo) cuando eran *mirados* por otro animal, posible pareja o depredador.

El psicólogo Julián Jaynes, en sus especulaciones atrevidamente originales de *Los orígenes de la conciencia en la ruptura de la mente bicameral*, enfatiza que la conciencia humana es una imposición cultural, extremadamente reciente, en una arquitectura funcional mucho más temprana.

### El efecto Balwin

Como sabemos, nada de lo que el individuo aprende es transmitido biológicamente a su proge-

nie. Sin embargo, los individuos suficientemente afortunados para estar cerca en diseño de un buen truco aprendible tenderán a tener más progenie, que también tenderá a estar cerca de ese buen truco. Si no fuera por la plasticidad -la capacidad de aprender-, este efecto, llamado efecto Balwin, no se presentaría. Gracias a él, la especie con plasticidad tenderá a evolucionar más rápido, y con más "dirección", que aquellas que no la tienen.

El genotipo humano incluye muchas adaptaciones que están ahí específicamente para favorecer la adquisición del lenguaje. Fueron adquiridas muy rápidamente, en términos de evolución, pero eso es precisamente lo que habría que esperar, dado el efecto Balwin: poder hablar es tan buen truco que cualquier individuo lento en adquirirlo habría estado en pavorosa desventaja.

### La autoexhortación

Llegó un tiempo en la evolución del lenguaje cuando las vocalizaciones empezaron a servir la función de compartir información útil. Un buen día un homínido, dotado de la capacidad de vocalizar, habría pedido ayuda -equivocadamente- cuando no había una audiencia servicial al alcance de su voz..., con excepción de sí mismo. Cuando habría oído su propia invocación, este estímulo habría provocado justamente la clase de ayuda que la invocación habría causado si hubiese venido de otro homínido: para su sorpresa, encontraría que el incidente le habría estimulado a contestar a su propia pregunta.

La práctica de hacerse preguntas a uno mismo puede ser claramente reconocida como un método de promover una acción mejor informada. Supongamos que, aunque la información correcta para algún propósito específico exista ya en el cerebro, estuviera en manos de un especialista incorrecto; el subsistema que necesitara la información no podría obtener la directamente -simplemente porque la evolución todavía no habría proveído la conexión interna adecuada. En estas circunstancias, podemos especular que las grandes ventajas de hablar *sotto voce* consigo mismo habrían sido fácilmente reconocidas, llevando pronto a lo homínidas a conversaciones internas completamente silenciosas.

Julian Jaynes ha argumentado persuasivamente que las capacidades del cerebro para autoexhortación y autorecordatorio son un prerre-

quisito de autocontrol sin los cuales la agricultura, los proyectos de construcción y otras actividades civilizadas y civilizadoras nunca hubieran podido organizarse. También ha postulado que un mecanismo arcaico para esta autoexhortación, basado en la división en dos "cámaras" cerebrales -correspondientes al invocador y al que responde a la invocación- podría haberse conservado hasta muy avanzados los tiempos históricos; tal bicamerismo, históricamente anterior a la autoconciencia, proveería una explicación, por ejemplo, a los relatos de apariciones de dioses en los tiempos mosaicos y homéricos.<sup>3</sup>

### Teoría de los memes

De acuerdo con R. Dawkins, las ideas se integran por ellas mismas en unidades memorables discretas, tales como las ideas de la rueda, de portar vestiduras, las ideas de venganza, del triángulo recto, del alfabeto, el calendario o la evolución por selección natural. En *El gene egoísta*, Dawkins nos urge a considerar seriamente la posibilidad de la evolución de estos racimos intelectuales o *memes* (por contraposición a *genes*). El autor argumenta muy persuasivamente que la teoría de la evolución por selección natural es neutral en relación con las diferencias entre memes y genes; ambos son solamente distintas clases de replicadores que evolucionan en diferentes medios a velocidades diferentes. Un carro con ruedas radiales no solo lleva grano o carga de un lugar a otro; lleva también la idea memorable de un carro con ruedas radiales de una mente a otra. La evolución de los memes no es simplemente análoga a la evolución biológica, sino un fenómeno que obedece *exactamente* las leyes de la selección natural como quedan especificadas por los requisitos de *variación, herencia y supervivencia diferencial*.

Dawkins sugiere que tendemos a pasar desapercibido el hecho fundamental de que "un aspecto cultural puede haber evolucionado en la forma en que lo ha hecho simplemente por ser ventajoso para sí mismo". No obstante, algunos memes son buenos también desde nuestra perspectiva, y no sólo desde la suya propia en calidad de autoreplicadores egoístas: los memes de la cooperación, la música, la escritura, la educación, la conciencia ambiental, la reducción de armamentos. Otros memes son más controversiales;

podemos entender por qué se extienden, y por qué debemos tolerarlos, a pesar de los problemas que causan: la burocracia, los embotellamientos de tráfico, la comida rápida, la propaganda política, el fútbol. Todavía otros más son incuestionablemente perniciosos y muy difíciles de erradicar: el antisemitismo, el secuestro de aviones, los virus informáticos, la corrupción administrativa. Los memes se expanden ahora alrededor del mundo a la velocidad de la luz, y se replican tan rápidamente que hacen aparecer la reproducción de la mosca de la fruta y la de las células de levadura como producidas en cámara lenta.

Puede llegar el día cuando evaluadores de memes no humanos basten para seleccionar y proveer a la preservación de memes particulares, pero por el momento, todos los memes tienen que pasar todavía una etapa "de capullo" en ese notable nido de memes que es la mente humana. Como las mentes están en corta provisión, y cada mente tiene una capacidad limitada para albergar memes, se produce considerable competencia entre ellos por entrar en cuantas más mentes sea posible. Esta competencia es la fuerza selectiva de la memosfera.

Por otro lado, las mentes humanas mismas son en grado considerable creación de los memes. Miles de memes, la mayoría sostenida por el lenguaje, pero algunas también por "imágenes" mudas y otras estructuras de datos, fijan cada día su residencia en algún cerebro individual y moldean sus tendencias, conformándolas en una mente.

Comparada con la plasticidad fenotípica, que ha estado con nosotros por millones de años, una evolución memética significativa es sólo un fenómeno extremadamente reciente: llega a ser fuerza poderosa en los últimos cien mil años y hace explosión vertiginosa a partir del desarrollo de la civilización hace menos de diez mil. Según sabemos, está restringida a una sola especie terrestre: *Homo sapiens*.

### La mente: máquina virtual del cerebro

La mejor manera de dar cuenta de la relación entre la biología básica del cerebro y su superestructura memética es usar una metáfora tomada de la informática. Podemos decir que las mentes conscientes humanas son *máquinas virtuales* más o menos seriales implantadas en un

*hardware* eminentemente paralelo suministrado por la evolución.

Dennett llega a la conclusión de que todo parece indicar que la conciencia humana

(1) es demasiado reciente para estar alambrada en nuestra maquinaria innata;

(2) es mayormente un producto de la evolución cultural que se imparte a los cerebros en el adiestramiento temprano; y

(3) su instalación exitosa está determinada por miríadas de micromodulaciones en la plasticidad del cerebro, lo cual significa que sus características funcionales son probablemente invisibles para el escrutinio neuroanatómico.

En nuestros cerebros existe una colección entreverada de circuitos especializados que, gracias a una familia de hábitos inculcados en parte por la cultura y en parte por autoexploración, conspira para producir una máquina virtual, más o menos ordenada, más o menos efectiva, más o menos bien diseñada, que podemos denominar *máquina joyceana*.<sup>4</sup> Esta máquina enyunta en una causa común a distintos órganos especializados con historias evolutivas independientes. Podemos considerarla como el *sistema operativo* o el *software* del cerebro, y es lo que realiza esa suerte de milagro político interior que consiste en crear en cada ocasión un *capataz virtual*, sin dar a ninguna parte del sistema poderes dictatoriales de largo plazo. ¿Quién está a cargo? Deberíamos decir que nadie, o que primero lo está una coalición de especialistas y después alguna otra, dependiendo de las circunstancias del trabajo interior.

### Una consecuencias desagradable de la virtualidad del "yo"

El hecho de que la conciencia, o "yo", no sea substancial (alambrado) sino virtual (programado y cambiante) parece tener valor explicativo sobre muchos aspectos de la realidad interior que han intrigado por igual a través de los tiempos a filósofos, moralistas, psicólogos y juristas. Algunos de esos aspectos, por ejemplo, tienen que ver con la patente crueldad que puede manifestar la agresividad humana, en todo disímil a la "crueldad necesaria" de la lucha por la supervivencia del mundo animal. Es bien posible que tales matices de encarnizamiento de la agresividad en los seres humanos<sup>5</sup> se base en la profunda inseguridad inducida en el "yo" por la conciencia de su propia

propia insubstancialidad: la mente es una máquina virtual que además *se sabe virtual*. No es ni siquiera el conocimiento de la propia mortalidad -asequible a algunos, si no a todos, los animales- lo que está en juego, sino el conocimiento de la propia *virtualidad*, del carácter eminentemente efímero de la conciencia como amalgama transitoria de volátiles coaliciones de sistemas especializados en el cerebro.

### Deficiencias y límites de la conciencia

El humilde origen de nuestra conciencia, como producto del *bricolage* de materiales preexistentes, queda claramente de manifiesto en sus abundantes límites y deficiencias. La chocante deficiencia en nuestra visión periférica (toda la visión excepto dos o tres grados alrededor del puro centro) se nos oculta normalmente por las trayectorias suaves o las repentinas sacudidas que hacen variar momentáneamente el contenido del área foveal central de la retina (el resto de la retina es, según se sabe, como diez veces menos discriminante que la fovea).

Otra característica sobresaliente de la conciencia es su discontinuidad, según se revela en el punto ciego de la retina y en las lagunas que dejan los movimientos de sacudida de los ojos, para tomar sólo los ejemplos más simples. La discontinuidad de la conciencia es llamativa en especial por lo frecuentemente que su continuidad ha sido alabada por profanos, literatos y filósofos <sup>6</sup>. Pero de hecho la conciencia es escasa y está llena de lagunas, no contiene ni la mitad de lo que la gente generalmente piensa que contiene.

Solemos maravillarnos, en momentos de autoconciencia intensificada, de la gloriosa riqueza de nuestra experiencia consciente. Pero esa riqueza de que nos vanagloriamos no es en efecto sino solamente la riqueza objetiva del mundo exterior, en su lujuriosa manifestación. Como lo hemos aprendido de innumerables experimentos en psicología cognoscitiva, tal riqueza nunca "entra" en nuestras mentes conscientes: simplemente permanece disponible para una conciencia activa pero intermitente, fragmentaria y limitada en muchos sentidos diferentes.

### La coevolución de la visión del color

Nuestra visión de color no nos da acceso a propiedades simples de los objetos, como lo pretendieran los filósofos empiristas, aun en este siglo. La visión en colores "naturales" coevolucionó *desde el principio* con colores cuya *raison d'être* era simplemente un *color coding* - como el que encontramos en las líneas de los trenes metropolitanos en algunos países- de rasgos naturales necesarios para la supervivencia. En esto parece que debemos volver las cosas "de cabeza": los arquitectos que diseñaron el metro de París, de México o de Washington D.C. no inventaron el *color coding*, sino que simplemente volvieron a utilizar los colores para su propósito primigenio. Algunas cosas en la naturaleza necesitaban ser vistas y otras necesitaban verlas, así pues evolucionó un sistema que tendió a minimizar la tarea de las últimas acentuando la perspicuidad de las primeras. El principio de la codificación por colores es la base de la visión de color en los insectos, no simplemente una reciente invención de una sagaz especie de mamífero. Los pájaros, los peces, los reptiles y los insectos tienen resueltamente visión en colores (en cambio, los perros y los gatos todavía no la han inventado: dentro de los mamíferos, solo los primates la han desarrollado).

### Las "clases naturales"

En el principio fueron los colores, para ser vistos por aquellos seres hechos para verlos. Una vez que hubo criaturas que distinguían las bayas rojas de las bayas verdes, pudieron ellas también distinguir los rubíes rojos de las esmeraldas verdes, pero sólo como una añadidura casual. El "cielo azul que todos vemos" es azul porque las manzanas maduras son rojas, las uvas moradas y la hierba verde y la vida necesitó como fondo de contraste el azul celeste.

Siguiendo el mismo argumento, no podemos preguntarnos con sentido si la diferencia entre las esmeraldas y los rubíes es una diferencia visible si no especificamos al mismo tiempo un sistema de visión determinado. Los colores no son "clases naturales" precisamente *porque* son el producto de la evolución *natural*, la cual tiene una tolerancia por fronteras desdibujadas entre

categorías que indignaría a los filósofos, aficionados a ideas claras y distintas. Dennett llega a decir que si la vida de una criatura dependiera de empaquetar juntos a la luna, el queso azul y las bicicletas, con toda seguridad la Madre Naturaleza encontraría una manera para hacerle ver estas cosas como "intuitivamente la misma y cabal clase de cosa".

Podemos concluir que no es objetivo de nuestros sistemas sensoriales el detectar propiedades "básicas" o "naturales" del entorno, sino sólo servir nuestro propósito narcisista de permanecer vivos; como dice Dennett, "no está entre las funciones de la naturaleza el crear máquinas epistémicas".

### La conciencia y la sociedad

Una de las deficiencias más reveladoras de la conciencia es su fragilidad. Según lo afirma Dennett, hay voluminosos diagnósticos perfectamente confiables de que no existen sólo un puñado o unos cientos sino miles de casos de DPM (desorden de personalidad múltiple) hasta el día de hoy. Casi invariablemente esos casos deben su existencia a ultraje prolongando en la primera infancia, usualmente de carácter sexual, o a severidad patológica en el seno de la familia. Así, se trata de un fenómeno muy complejo que se extiende mucho más allá de los cerebros individuales de quienes lo sufren. Aquí, como en muchas otras ocasiones, una deficiencia de la conciencia nos pone sobre la pista de una de sus notas fundamentales; en este caso su carácter eminentemente social. Se puede considerar como establecido por la ciencia que las personalidades no son madreperlas que existan independientemente, sino productos de los procesos sociales que nos crean; como tales artefactos, están sujetas a cambios repentinos de estatus. La única estabilidad que podemos atribuir a la trayectoria de una personalidad, como a la de un club, es el *momentum* que le imparte la red de creencias que la constituye; y cuando estas creencias se reblandecen, aquella falla, sea temporal o permanentemente.

Como lo ha mostrado Marvin Minsky en *La sociedad de la mente*, existe una continuidad natural entre la dinámica que origina y explica el funcionamiento de las mentes y la que origina y explica el funcionamiento de las sociedades: ambas resultan de la interacción de innumerables

procesos paralelos que se interpenetran en todas direcciones. Así pues, una personalidad no es un punto matemático ni un ente metafísico, sino una abstracción definida sobre las miríadas de atribuciones e interpretaciones (incluidas las autoatribuciones y las autointerpretaciones) que componen la biografía de un cuerpo viviente cuyo *centro de gravedad narrativa* (como diría Dennett) es ella misma. Como tal, juega un papel singularmente importante en la economía cognoscitiva del cuerpo viviente, porque de todos los modelos de cosas del entorno que un cerebro activo debe construir mentalmente ninguna es más crucial para su supervivencia y bienestar que el modelo que el agente se hace de sí mismo.

### Referencias

- Churchland P.S., *Neurophilosophy*. (Cambridge, MA: MIT Press, 1986).
- Dennett D.C., *Consciousness Explained*. (Boston: Little, Brown and Co., 1991).
- Dawkins R., *The Selfish Gene*. (Oxford: Oxford University Press, 1976).
- Dawkins R., *The Blind Watchmaker*. (New York: Norton, 1986).
- Fischler M.A. & O. Firschein, *Intelligence: The Eye, the Brain, and the Computer*. (Reading, MA: Addison-Wesley, 1987).
- Gazzaniga M., *The Social Brain*. (New York: Basic Books, 1985).
- Jackendoff R., *Consciousness and the Computational Mind*. (Cambridge, MA: MIT Press, 1987).
- Jaynes J., *The Origin of Consciousness in the Breakdown of the Bicameral Mind*. (Boston: Houghton Mifflin, 1976).
- Kolers P.A. & M. von Grünau, "Shape and Color in Apparent Motion", *Vision Research*, **16**, 1976.
- Gutiérrez, C. *Nueve ensayos epistemológicos* (San José, C.R.: Editorial Costa Rica, 1982).
- Goodman N., *Ways of Worldmaking*. (Hassocks, Sussex: Harvester, 1978).
- Minsky M., *The Society of Mind*. (New York: Simon and Schuster, 1985).

Quine W.V.O, "Natural Kinds," *Ontological Relativity and Other Essays*. (New York: Columbia University Press, 1969).

## Notas

1. Un arco reflejo sería, por ejemplo, un tal *átomo de inteligencia*.
2. Se llama ingeniería inversa al proceso de descubrir el diseño a partir de la máquina. Bien podríamos considerar a

la ciencia toda como un proyecto de ingeniería inversa para discernir el diseño del universo.

3. La autoconciencia misma, según Jaynes, no habría comenzado a existir sino a partir de los tiempos de Sócrates y Pericles.

4. Dennett le da este nombre en honor del célebre novelista irlandés James Joyce, que introdujo en la literatura el "flujo de conciencia".

5. Debo la percepción de este aspecto de la vida mental a una sagaz intervención de Andrés Sáenz con ocasión de la presentación oral de estas ideas en la Universidad Nacional.

6. Por ejemplo por Kant o Bergson.

"Venus no es pero la estatua existe, (...) si no ha seguido siendo piedra en las montañas de Carrara, es porque se le ha dado la función de figurar un cierto no-ser; pero inversamente, puesto que ese no-ser como tal es reconocido como consecuencia de la imaginario social, todo el objeto es ventilado en su ser; la sociedad reconoce pues una verdad ontológica en la medida en que el ser de ese objeto es considerado como inclinación permanente a desrealizarse al irrealizarse a su material en Venus. El ser compuesto a partir de lo real para desrealizarse públicamente al darse a conocer en el mundo".

Claudio Gutiérrez  
University of Delaware  
Newark, Del. 19716  
USA.

Algo análogo puede afirmarse de todas las artes. En el cine y el teatro "el actor se presenta bajo la personalidad de un personaje *irreal*, en la pintura la perspectiva es engañosa; la novela nos introduce en un mundo que sólo está constituido por imágenes verbales (...). Esto equivale a decir (...) que la esencia del objeto estético es su *irrealidad*", como es también el caso en los sueños y trastornos de la mente, aunque allí las imágenes son más efímeras que una página impresa y la tela pintada.

El mundo onírico es un medio imaginario. Freud decía que "el sueño es la realización (*disfrazada*) de un deseo reprimido". Tal idea transita en las interpretaciones contemporáneas de la vida onírica, aunque matizada por las diversas escuelas de psicología profunda. Pero allí la necesidad de tenuizar el campo trascendental de la conciencia en la vigilia y el sueño ha sido apenas intuida. Es Dieter Wyss (fenomenología amplificadora), en sus investigaciones oniológico antropológicas, quien se ha ocupado más sistemáticamente de tal problemática. En obras como *¿Conciencia onírica?* y *La filosofía del caos o lo irracional*, asigna a la imaginación un papel constitutivo-constituyente en el flujo vivencial, cuando distingue las funciones personalizantes de la

*Summary: The author considers the imagination in dreams, art and madness, and views the possibility that every world aperception implies constitutive imaginary synthesis, which ontological modality can be inquired through the description of the noetic-noematic structure of the intentionality.*

*Resumen: El autor considera la imaginación en el sueño, el arte y la locura, y concibe la posibilidad de que toda apercepción del mundo implique síntesis imaginarias constituyentes, cuya modalidad onto-lógica puede indagarse describiendo la estructura noético-noemática de la intencionalidad.*

## 1. Correlatos Imaginarios

Hay seres reconocibles principalmente por la función imaginaria que desempeñan. Su cantidad y variedad es indefinida, pues el poder de imaginar se objetiva en formas muy diversas; uno tanto para inventar modelos explicativos de la realidad como para representar el rostro de una persona ausente o crear universos fantásticos, desde los árboles hablan con los hombres, las aves tienen alas de mariposa y las máquinas pueden sentir pasiones humanas.

"Una estatua es una mujer imaginaria", ha escrito Sartre a propósito de un trozo de mármol que representa a Venus. Su materialidad es el soporte permanente de alguna *irrealización*, pero la estatua de *irrealidad* -propio de las imágenes- es necesario para que la materia sea objeto estética. Si el mármol *representa* a Venus es porque su ser está determinado en lo imaginario, pero figura eso en la siguiente forma: