

FUNCIONALISMO MENTAL Y TEORIA NEUROBIOLÓGICA DE LA CONSCIENCIA

Camilo J. Cela Conde

Departamento de Filosofía. Universidad de las Islas Baleares

Durante los días 5, 6 y 7 de mayo de 1993 el profesor Francis Crick permaneció en la Universidad de las Islas Baleares, invitado por el Departamento de Filosofía y dentro del programa de celebración del decimoquinto aniversario de la U.I.B. El profesor Crick pronunció una conferencia acerca de la Teoría neurobiológica de la consciencia, que es el aspecto más relevante de sus investigaciones desde que se incorporó al grupo PDP (Parallel distributed processes) en el Salk Institute de la Universidad de California (San Diego). Sirvan estas notas para centrar algunos aspectos de esa teoría y su papel histórico como alternativa al funcionalismo mental imperante en la Psicología cognitiva de los años setenta y ochenta.

Palabras clave: cerebro, mente, consciencia, neurobiología, funcionalismo, visión.

Mental functionalism and neurobiological theory of consciousness. On May 1993, Professor Francis Crick was invited by the Department of Philosophy (University of Balearic Islands, at Palma de Mallorca) to give a lesson on his Neurobiological Theory of Consciousness -that is the most relevant aspect of Professor Crick's contribution to the PDP (Parallel Distributed Processes) Group. May these notes be useful to focus some aspects on that theory and its historical role as an alternative to the mental functionalism of most cognitive psychologists along the 70' and 80' years.

Key words: brain, mind, consciousness, neurobiology, functionalism, visual awareness.

El auge de la psicología cognitiva durante la segunda mitad de nuestro siglo supone la puesta en marcha de un estudio de los procesos mentales que se plantea, en líneas generales, como una reacción contra los modelos de «caja negra» de explicación de la conducta utilizados por los conductistas. Y el objeto declarado de la nueva forma de hacer psicología es el de establecer de qué forma se genera y almacena el conocimiento en la mente humana. Ese plan de trabajo implica un cambio radical en los planteamientos teóricos y metodológicos. Objetivos de estudio abandonados por el conductismo, como la consciencia, la intención o la voluntad, se convierten en ejes del nuevo paradigma.

Pero la sensación de bloque que, como alternativa a los conductistas, presenta el cognitivismo no puede hacernos olvidar que en sus inicios se apuntaban ya dos soluciones muy distintas al problema del conocimiento. La primera, ligada a los trabajos de Rosenblatt, fue pronto abandonada en favor de la segunda escuela, la que suele conocerse con el nombre de metáfora del ordenador y supone, por lo que hace a sus planteamientos, una continuidad de las ideas de Von Neumann y Turing¹. Fue la metáfora de la

¹ Bunge (1988) sostiene que las metáforas no son teorías; según él, son meras descripciones fantásticas sin valor explicativo. Parece evidente, sin embargo, que la metáfora de la computadora no es una mera descripción fantástica y *pretende* tener valor explicativo. El propósito de Turing no es en absoluto el de hacer fantasías, y así parece haberse entendido por parte tanto de sus seguidores como de sus numerosos críticos.

computadora, pues, el eje del desarrollo de las «ciencias cognitivas» en sus dos líneas de avance: las investigaciones acerca de la Inteligencia Artificial y los modelos de la mente funcionalista.

Sostener que el cerebro funciona como una computadora tiene la ventaja de poder estudiar cosas tan complejas como son el pensamiento, la mente, la atención, la consciencia, etc., dando por supuesto que obedecen a ciertas reglas extraídas del examen, mucho más sencillo, de la forma de operar del ordenador. Los rasgos más sobresalientes de la metáfora de la computadora aplicada al conocimiento humano son: 1) el soporte biológico (el equivalente al hardware) y los programas de procesamiento (el software) son distinguibles, irreductibles entre sí y explicables totalmente por separado²; 2) existe un sistema de procesamiento central que recoge y transforma los datos procedentes del mundo exterior; 3) la información se procesa de forma serial y 4) la información se almacena de dos formas distintas; es decir, fugazmente (memoria de trabajo) y a largo plazo (memoria permanente).

Cierto es que, desde muy pronto, no faltan voces que advierten de lo inoportuno de la comparación metafórica entre cerebro y computadora. Un caso notorio es el de Eccles (1976), quien, mediante un barrido de los estudios de Hubel, Wiesel, Colonnier y Szentágothai, ofrece una idea bastante precisa de los rasgos microscópicos de la estructura cerebral humana y, concretamente, del córtex. Aunque ésta varía según las zonas examinadas, pueden indicarse ciertas características generales:

- aunque existen más de 40 áreas en el córtex funcionalmente distintas (áreas de

Brodmann), todas las partes del córtex muestran una disposición básica de su estructura vertical en capas (normalmente seis)

- la respuesta a diferentes estímulos sensoriales tiene lugar, en cada área, mediante la intervención de distintos tipos de neuronas especializadas y agrupadas en zonas corticales en forma de columna (la forma difiere en distintas regiones del neocórtex, pero mantiene varias similitudes que permiten hablar, en general, de módulo columnar)

- esas agrupaciones de córtex cerebral en forma de columna de neuronas son las unidades básicas, los «módulos», desde el punto de vista no solamente estructural sino también funcional. Constituyen unidades de procesamiento de información, en los términos usuales de la época.

El «módulo» (Szentágothai) es el eje de un circuito neuronal básico constituido por canales de entrada de impulsos (fibras aferentes), interacciones neuronales complejas (conexiones dendríticas entre células piramidales, diversos tipos de interneuronas excitadoras -neuronas estrelladas- y neuronas inhibitorias de tipo cesta, con un total de hasta 10.000 neuronas implicadas) y canales de salida de impulsos (los axones de las células piramidales). Pero al margen de los detalles anatómicos, lo importante del modelo modular del córtex es que las operaciones que realiza un módulo cualquiera pueden ser imaginadas como un complejo de circuitos en paralelo (con la suma de centenares de líneas de convergencia entre neuronas) al que se añade un enredo de conexiones en vaivén entre líneas excitantes e inhibitorias. Como cada uno de los módulos puede actuar sobre centenares de otros módulos, recibiendo a su vez la acción de otros tantos, la complejidad de la operación de todo el agregado de unos dos millones de módulos está, en palabras de Eccles, más allá de todo lo imaginable.

Algo sí puede, sin embargo, imaginarse de inmediato: la manera de funcionar de la computadora y del cerebro es radicalmente

² Este principio impone de por sí ciertas respuestas al problema de las relaciones existentes entre mente y cerebro. Más tarde se hará referencia a él; mientras tanto, «mente» y «cerebro» se entenderán como conceptos relativos, respectivamente, a las funciones y a la estructura cerebral.

distinta. El funcionamiento serial de la computadora contrasta con el funcionamiento en paralelo del cerebro. Las velocidades de transmisión y el número de unidades de procesamiento de la información son, en uno y otro caso, muy distintas. No es raro que en el trabajo de Eccles encontremos una clara advertencia:

«tenemos que enfatizar que los niveles de complejidad en la operación de un módulo van mucho más allá de los microcircuitos electrónicos y deben ser considerados como de un orden totalmente distinto» (1976, p. 109).

Pero el énfasis en las explicaciones funcionales y el éxito de los modelos del lenguaje y la mente de autores como Chomsky, Fodor o Johnson-Laird se encargaron, en esos años, de minimizar las diferencias. En la psicología cognitiva de entonces sigue teniendo una gran importancia el modelo funcional de la mente.

EL MODELO FUNCIONAL DE LA MENTE

Desde el punto de vista funcional, la metáfora de la computadora proporciona un modelo de procesamiento de la información en el que un rasgo sobresale sobre los demás al hablar de la arquitectura de la mente: existe un «procesador central» que se encarga de manipular las señales aportadas desde fuera.

Para hacer compatible esa idea con las limitaciones orgánicas más obvias del ser humano Fodor (1983) propone un modelo de la forma como se procesan las informaciones sensoriales, modelo que contiene dos filtros. Un primer filtro lo forman los transductores, que convierten la estimulación proximal de los sentidos en señales neuronales. Tal como dice Fodor, «la percepción está destinada a representar el mundo de modo que éste sea accesible al pensamiento» (1983, p. 67). El segundo filtro lo constituyen los sistemas de entrada: unos ele-

mentos que median entre los productos de los transductores y los mecanismos cognitivos centrales, es decir, que «efectúan la codificación de las representaciones mentales que constituyen el ámbito de operación de los mecanismos centrales» (Fodor, 1983, p. 69). El papel de los sistemas de entrada es el de transformar los diferentes formatos representacionales de los distintos sentidos en un formato único que es el apropiado para la computación realizada por el sistema central³.

El uso de la metáfora de la computadora no es exclusivo de Fodor: está presente, con diversos matices, en todos los funcionalistas. Así Johnson-Laird (1988), en una obra considerada como uno de los hitos del funcionalismo, incluye declaraciones de principios como las que siguen:

«La respuesta al problema no es directamente el ordenador, sino lo que descansa en él, su predecesor matemático, la teoría de la computabilidad. Esta teoría, que fue inventada por los lógicos de los años treinta, (...) muestra cómo puede utilizarse un conjunto elemental de unidades constituyentes para construir una ilimitada variedad de procesos simbólicos complejos (...). Hasta ahora, ha demostrado ser aplicable a cualquiera dominio y es particularmente apropiada para modelar ese otro dispositivo que parece ser capaz de manipular símbolos: el cerebro humano.» (1988, pp. 25-26).

«El cerebro humano no puede contener un número infinito de símbolos preexistentes (...). La vasta variedad de símbolos mentales debe construirse mediante medios finitos, mediante símbolos primitivos. Los impulsos nerviosos y los demás acontecimientos electroquímicos pueden considerarse, pues, como los primitivos subyacentes -

³ Un modelo funcional así implica, al menos, dos consecuencias importantes: 1. No existen las «imágenes mentales» (Pylyshyn, 1984) 2. Las representaciones mentales tienen un mero sentido sintáctico: la semántica no tiene en él lugar (Fodor, 1983).

quizá análogos en forma- mediante los que se construyen los símbolos. Aquí el ordenador y el poder de los símbolos numéricos proporcionan la idea crucial. Al margen de lo complicada que sea la actuación del ordenador, de cuál sea el dominio, o de cómo sean de profundos los resultados, todo lo que puede hacer un ordenador es llevar a cabo unos pocos tipos de operaciones básicas sobre cifras binarias. En consecuencia, los procesos computacionales deberían ser capaces de modelar los símbolos mentales y sus manipulaciones». (1988, p. 35).

Johnson-Laird mantiene en su teoría de la mente aspectos importantes de la metáfora de la computadora, como son la existencia de un procesador central, la manipulación de símbolos simples en términos sintácticos o la distinción entre *software* y *hardware*. Pero también incorpora algunas modificaciones, como es el hecho de admitir que la mente procesa la información en paralelo, y no en serie. Una idea que había introducido, dos años antes, el grupo PDP a través de una alternativa radical a los modelos de la metáfora de la computadora: el del «procesamiento paralelo distribuido» por parte del cerebro⁴.

QUIEBRA DE LA METAFORA DE LA COMPUTADORA

El modelo de la mente propuesto por el grupo PDP (Rumelhart y McClelland, 1986) Incluye varios rasgos opuestos a la metáfora de la computadora. En primer lugar, la arquitectura mental no presupone la existencia de reglas explícitas de procesamiento: el *software* y el *hardware* dejan de ser referen-

cias válidas. Además, la memoria no tiene una localización precisa: se distribuye a lo largo de varias unidades de procesamiento. Es evidente también, por el propio nombre de la escuela, que el procesamiento de la información se realiza en paralelo –tal como apuntaban los neurobiólogos– y no serialmente.

Se derivan ya de ese planteamiento algunas consecuencias importantes como, por ejemplo, el hecho de que el aprendizaje de una lengua no es la adquisición de la gramática en sentido chomskiano. Pero los aspectos más importantes son los del vuelco de la metáfora: no solamente el cerebro no funciona como una computadora sino que toda la estrategia de la Inteligencia Artificial se pone en evidencia y se vuelve a las máquinas como el «perceptron» (Rosenblat, 1961), es decir, máquinas que pretenden realizar tareas inteligentes incorporando, en cierto modo, las operaciones cerebrales⁵. Pero es la incorporación de Francis Crick al grupo PDP la que provoca la aparición de un énfasis en las explicaciones estructurales de la actividad cerebral y la propuesta de una de las teorías más interesantes para poder establecer la relación existente entre mente y cerebro. Conjuntamente con Christof Koch (del California Institute of Technology), Crick ha esbozado lo que puede ser una teoría neurobiológica de la consciencia (Crick y Koch, 1990).

En una entrevista concedida a Jubak, Koch admite que la teoría neurobiológica de la consciencia tiene una finalidad provocativa: la de intentar que los investigadores de redes neuronales y los neurobiólogos comiencen a cambiar las preguntas que se hacen. La consciencia, ciertamente, no figu-

⁴ «Procesamiento paralelo distribuido», «conexionismo» y «sistema de red neural» son formas distintas de calificar estos modelos. Como se dijo al principio, los procesos en paralelo eran algo ya propuesto en los inicios del cognitivismo (McCulloch, Pitts, Rosenblat), pero se habían desechado por la explicación entonces más clara y convincente de la metáfora de la computadora.

⁵ Changeux entiende que la construcción de máquinas «neuromiméticas» es un tercer paso después de la Inteligencia Artificial basada en la metáfora de la computadora y los modelos conexionistas (1989, pp. 203-204). Para una descripción divulgativa de algunas de esas máquinas y sus paralelos con el funcionamiento del cerebro, vid. Jubak (1991).

raba hasta entonces en el catálogo de los conceptos que utilizaban ni unos ni otros. Pero, por otra parte, la coherencia de la postura de Crick y Koch es absoluta: si se mantiene que las funciones y la estructura del cerebro están estrecha e indisolublemente relacionadas, para poder explicar los fenómenos como atención, consciencia, memoria, intención, etc. —centrales dentro del paradigma cognitivo— habrá que entender cuáles son las bases neuronales que los hacen posibles.

El punto de partida de la teoría de Crick y Koch es una sugerencia de Von der Malsburg acerca de la sincronización de neuronas distantes en el cerebro como elemento capaz de relacionarlas con un determinado objeto (Von der Malsburg, 1973)⁶. Frente a las actividades azarosas de neuronas —coincidentes de manera aleatoria— una activación rítmica, con un ritmo compartido, podía ser el elemento que definía la pertenencia de diversas neuronas al conjunto activado por un comportamiento sensorial. Siguiendo esa idea, el grupo de Charles Gray identificó respuestas neuronales sincronizadas a las estimulaciones de la retina del gato (Gray, Raether y Singer, 1989; Gray, Engel, König y Singer, 1990) lo que, literalmente, suponía la primera identificación estructural de un «pensamiento»⁷. Pero han sido Crick y Koch (1990) los que han extendido esas evidencias experimentales hacia una teoría amplia capaz de relacionar actividad neuronal, atención y consciencia.

TEORIA NEUROBIOLÓGICA DE LA CONSCIENCIA

La teoría neurobiológica de la consciencia de Crick y Koch establece que ésta de-

pende, en primer lugar, de una cierta memoria a corto plazo a la que se añade cierto mecanismo de serial de atención. Para ello es preciso suponer dos cosas:

1. que la consciencia, al menos parcialmente, tiene que ver con la actividad conjunta de distintas neuronas

2. que los diferentes tipos de consciencia (visual, auditiva, etc.) utilizan un mismo mecanismo básico.

El primer supuesto se remota a Hebb (1949), pero aprovecha los hallazgos experimentales del grupo de Gray. El segundo permite que los fenómenos se estudien en función de las actividades sensoriales más simples y fáciles de experimentar, para generalizar después como modelo general de la consciencia. Si se admiten esas dos presunciones, se puede hacer un esbozo estructural de cómo y dónde actúan los correlatos neuronales de la consciencia:

— la consciencia correlaciona con una especial actividad (cuya forma se verá luego) de un subconjunto de neuronas del sistema cortical

— las operaciones ocurren, principalmente, en el neocórtex y las estructuras asociadas (tálamo, ganglios basales y claustró), es decir, en lo que Crick y Koch llaman el «sistema cortical»

— el paleocórtex, asociado con el sistema olfativo, también está probablemente implicado

— el allocórtex (hipocampo) no es esencial, pero puede intervenir

— el cerebro medio y profundo (cerebelo) no son esenciales.

Ese modelo elimina ya algunos supuestos anteriores de la ciencia cognitiva. En la medida en que el subconjunto que se activa implica neuronas muy alejadas unas de otras, Crick y Koch despachan, de forma un tanto irónica, la vieja y querida idea del funcionalismo de la existencia de un procesador central. «Si existe un sistema operativo de ese tipo no es fácil, por el momento, ver ningún área particular del

⁶ La idea es antigua. Fue Donald Hebb (1949) quien primero sugirió que la actividad conjunta de un grupo de neuronas es la clave de las representaciones mentales.

⁷ Vid. la entrevista de Jubak a Gray (Jubak, 1991, cap. 14).

cerebro en la que esté localizado» (Crick y Koch, 1990, p. 265)⁸.

La activación conjunta de un subgrupo de neuronas corticales desecha la idea de correlación biunívoca entre el objeto visual y una única neurona -la teoría de la «neurona abuela». En cada momento, las neuronas relevantes de numerosas áreas corticales cooperan conjuntamente para producir la consciencia visual.

Ese fenómeno destruye también, aunque Crick y Koch no lo mencionan explícitamente, otro de los modelos centrales del cognitivismo inicial: el de la memoria multialmacén (Atkinson y Shiffrin, 1968). La presencia de distintos almacenes para almacenar recuerdos a corto y largo plazo, construido, como no, gracias a la metáfora de la computadora, choca frontalmente con la representación mental como actividad conjunta de neuronas. De igual forma que no hay un procesador central operativo que pueda localizarse en el cerebro, tampoco parece haber un lugar específico destinado a almacenar recuerdos. Estos consisten en lazos establecidos entre neuronas de una forma más o menos profunda -y por tanto, más o menos permanente-; algo así como un «paisaje neuronal»⁹, en el que el subgrupo de neuronas que se activa simultáneamente deja huellas de su presencia y permite posteriores activaciones más o menos fieles a la inicial. Es importante tenerlo en cuenta porque Crick y Koch utilizan la terminología de «memoria episódica», «memoria a corto plazo» y «memoria a largo plazo» propia de los modelos multialmacén. Sería una equivocación el

creer que esos conceptos implican, en la teoría neurobiológica de la consciencia, la presencia de almacenes separados e identificables.

COMPONENTES ESTRUCTURALES DE LOS LAZOS NEURONALES

La teoría de Crick y Koch es, como se ha dicho, una generalización de las evidencias experimentales obtenidas del estudio del sistema visual (principalmente el del macaco). Los autores indican los motivos por los que se elige ese sentido en particular: lo comparan tanto los seres humanos como los primates, existe mucho trabajo experimental en ese terreno y el tipo de consciencia implicado en él es simple¹⁰. Pero a través del supuesto de identidad, el modelo obtenido en base a las representaciones visuales puede extenderse a la consciencia en general.

El sistema visual de los mamíferos es complejo, con diferentes áreas visuales en el neocórtex (más de veinte en el macaco -cercopitecoideo). Las diferentes áreas corticales corresponden, en general, a distintos rasgos simples del campo visual (forma, color, movimiento, orientación) y la consciencia visual es, como establece el modelo de Crick y Koch, el resultado de la actividad global de las diversas áreas. Eso supone que, en un determinado momento, las neuronas relevantes de una multitud de áreas corticales deben cooperar conjuntamente: cada objeto en el campo visual es representado por la activación de un conjunto de neuronas que se encuentran distribuidas entre diferentes áreas corticales.

¿Cómo se vuelven activas las neuronas de un conjunto tan amplio y tan disperso que, en casos como el de reconocimiento de un rostro, debe implicar incluso las huellas de la memoria a largo plazo? Como ya había indi-

⁸ La «Primera Ley Fodoriana de inexistencia de la Psicología Cognitiva», por la que Fodor establecía la imposibilidad de estudiar la actividad del procesador central, recibe así una irónica confirmación: desde luego será imposible estudiar algo que no existe. Para una crítica de esa ley vid. Cela Conde y Marty (1991).

⁹ Similar, en cierto modo, al «paisaje epigenético» de Waddington (1975), que ha sido recuperado y ampliado a través de la teoría de las reglas epigenéticas por Lumsden y Wilson (1981, pp. 35 y ss.) Vid también Changeux (1989).

¹⁰ Para la distinción entre la consciencia simple y autoconsciencia o consciencia compleja vid. Johnson-Laird (1988) p. 355.

cado Von der Malsburg, es necesario que exista un lazo, algo común. Ese lazo es el mecanismo capaz de señalar las neuronas que generan, al activarse de forma correlacionada, la percepción de esa determinada persona. Pero el modelo, por ahora, continúa siendo de caja negra. El aporte esencial de la teoría neurobiológica de la consciencia procede de dos propuestas de Crick y Koch relativas, respectivamente, a la estructura de los lazos neuronales y al papel de la atención en el proceso.

Vayamos con la primera. El modelo estructural de lazos neuronales de Crick y Koch (de carácter hipotético, excepto por lo que hace al primer punto) es como sigue:

1. Los lazos consisten en la oscilación semisincronizada de las neuronas en la gama de los 40-70 Hertzios.
2. El cerebro tiene una especie de mapa de relieve topográfico (puede que en el tálamo) en el que los objetos del entorno visual se encuentran enfatizados por sus características más sobresalientes.
3. Uno de los «picos» se selecciona mediante un mecanismo de todo-o-nada.
4. Las características de ese tipo se relacionan con las características del mapa visual perceptivo de rasgos, en un proceso de feed-back.
5. Ese proceso de feed-back es el que produce las oscilaciones y, por tanto, la consciencia.

CONSCIENCIA Y ORIENTACION ESPACIAL

El subgrupo de neuronas de la zona cortical que correlaciona con la consciencia visual de un objeto no tiene orientación espacial en las distintas áreas del cerebro; estamos muy lejos de la idea del homínulo, del espectador sentado dentro del cerebro que se suponía que contemplaba, como en una pantalla, representaciones tridimensionales del mundo exterior. Pero Changeux (1989), de la mano de las investigaciones de

Georgopoulos (Georgopoulos, Schwartz y Kettner, 1986), indica cómo en el cerebro del macaco aparece una clara organización espacial. Examinando la actividad individual de unos centenares de neuronas del córtex motor cuando el animal indica, con su mano, una dirección en el espacio, Georgopoulos ha establecido el correlato de esa dirección con la actividad neuronal. Las neuronas son más o menos activas según cuál sea la dirección indicada por el mono: para cada neurona hay una dirección determinada, la «dirección óptima», en la que la actividad es máxima. De tal forma, la respuesta de las neuronas puede representarse mediante un vector cuya dirección es la de la dirección óptima, y cuya longitud es proporcional a la actividad de la neurona. Esa longitud, obviamente, es máxima cuando el animal está señalando en la dirección que corresponde a la óptima de la neurona en particular.

Tomando el conjunto de las neuronas, definiendo sus vectores y llevando a cabo su suma algebraica, se obtiene un vector resultante que coincide, con menos del 10% de error, con la dirección que está indicando el animal en un momento determinado. Eso quiere decir que a la dirección macroscópica indicada por la mano le corresponde, muy estrechamente, una actividad microscópica en forma de suma de vectores del conjunto de neuronas implicado. Changeux cree que ese modelo es general y no específico del caso experimentado por Georgopoulos. Siempre que se dé un cierto nivel de complejidad en el sistema nervioso central, se organizan «objetos mentales» de esa forma, mediante el estado de actividad y la «gráfica» de un conjunto de neuronas (Changeux, 1989, p. 136 y ss.).

LAZOS NEURONALES Y ATENCION

Los trabajos de Georgopoulos indican que la activación conjunta de neuronas no es exclusiva de la consciencia visual: las neuronas pueden establecer lazos de ese tipo que

tengan que ver con otro tipo de fenómenos de representación mental. Damasio (1989), por su parte, describe dos tipos de lazos:

– epigenéticos (fijados por selección natural y controlados genéticamente, como los que percepción espacio-temporal; probablemente son de este tipo los descritos por Georgopoulos)

– adquiridos (lazos muy fuertes establecidos a través de una experiencia continuada, como los del reconocimiento de los caracteres de un alfabeto)

Para los efectos de la teoría neurobiológica de la consciencia, Crick y Koch añaden un tercer tipo: los lazos esporádicos (lazos muy rápidos, prácticamente ilimitados y transitorios, que sirven para el reconocimiento de objetos novedosos). Estos últimos lazos son los correspondientes a la consciencia visual. Y el hecho de que se establezcan (junto con la manera como se establecen) tiene que ver, precisamente, con el segundo de los grandes aspectos de la teoría de Crick y Koch: el papel jugado en la consciencia por el mecanismo serial de la atención (Wolfe, Cave y Franzel, 1989).

El que la consciencia visual es el resultado de la atención fue sugerido ya por William James un siglo atrás. La novedad consiste en la propuesta de un modelo funcional acerca de la manera como la atención provoca la consciencia. Según Crick y Koch, el cerebro dispone de una especie de mapa topográfico de relieves en el que están representadas las localizaciones del campo visual de forma más o menos conspicua. Es algo así como una derivación del mapa de rasgos individuales, muy sesgada, donde están enfatizadas aquellas localizaciones en las que los objetos difieren (en color, movimiento, profundidad, etc.) de los que les rodean. Cuando la atención selecciona una localización particular, se produce un movimiento en feed-back entre el mapa de relieves y el de rasgos individuales.

Para que se produzca la consciencia de la forma indicada, el cerebro ha de decidir, uti-

lizando el conocimiento categórico almacenado en las conexiones, cuáles son las neuronas que han de oscilar conjuntamente para producir una representación verídica del objeto al que se atiende. Crick y Koch confiesan que está por construirse una teoría cognitiva acerca de cómo lo hace. Pero sugieren algunas posibilidades. Así, la distinción entre diferentes tipos de consciencia. Es posible que exista, primeramente, una «consciencia fugaz» (*fleeting awareness*), conectada a la memoria icónica, con una gama muy amplia de objetos contenidos en ella de forma transitoria y que permite dotar de una gran riqueza el campo visual.

Posteriormente, la atención se enfoca en una parte de esta consciencia fugaz, apareciendo la «consciencia de trabajo» (*working awareness*). Los objetos cuyas representaciones neuronales han alcanzado este nivel de consciencia son depositados en la memoria de trabajo y las propiedades asociadas a su localización se recuerdan durante un plazo corto de tiempo. Está por establecer de qué forma otras formas de memoria (a largo plazo) intervienen en el proceso¹¹.

La teoría de la consciencia de Crick y Koch deja, por supuesto, numerosos cabos sueltos. Ellos mismos reconocen que contiene probablemente graves errores que más tarde, cuando la experimentación avance, resultarán obvios¹². Pero existen numerosas dificultades para establecer investigaciones experimentales en ese terreno, dificultades denunciadas hace poco por Crick. La teoría actual se basa, como hemos visto, en el conocimiento del córtex visual de macaco; para poder avanzar en el estudio de la consciencia es necesario contar con unas herramientas que permitan asegurar cómo es la estructura en el ser humano. Las actuales o son muy limitadas

¹¹ Otro aspecto pendiente es el de la integración de señales procedentes de la actividad emotiva en la producción de la consciencia.

¹² Koch, en conversación con Jubak (Jubak, 1991).

(como es el caso de la resonancia magnética, de tipo estático, o la tomografía por emisión de positrones, dinámica pero con una resolución que, en la mayoría de los casos, alcanza 7-8 mm.-un mm³ del córtex alberga 40.000 neuronas) o resultan simplemente inadmisibles (implican trepanación y manipulación del cerebro). Así que los verdaderos avances no se podrán lograr mientras no aparezcan nuevos medios capaces de superar las barreras de tipo técnico o ético que limitan los experimentos de ahora mismo (Crick y Jones, 1993).

REFERENCIAS

- Ayala, J. F. (1970). «Teleological explanations in evolutionary biology». *Philosophy of Science*, 37, 1-15.
- Cela Conde, C. J. y Marty, G. (1991). «El ciudadano Fodor no existe. Estructura y función en los procesos cognitivos». *Psicotema*, 3, 231-241.
- Changeux, J. P. y Connes, A. (1989). *Matière à pensée*. Paris, Odile Jacobs.
- Crick, F. y Koch, C. (1990). «Towards a neurobiological theory of consciousness». *Seminars in the Neurosciences*, 2, 263-275.
- Crick, F. y Jones, E. (1993). «Backwardness of human neuroanatomy». *Nature*, 361, 109-110.
- Damasio, A. R. (1989). «The brain binds entities and events by multiregional activation from convergence zones». *Neural Computation*, 1, 123-132.
- Eccles, J. (1976). «Brain and Free Will». En Globus, G.G., Maxwell, G. y Savodnik, I. (eds.), *Consciousness and the brain. A Scientific and Philosophical Inquiry* (pp. 101-121). New York, N.Y., Plenum Press.
- Fodor, J. A. (1975). *The Language of Thought*. New York, N.Y., Crowell.
- Fodor, J. A. (1983). *The Modularity of Mind*. Cambridge, Mass., M.I.T. Press. Ed. española, *La modularidad de la mente*, Madrid, Morata, 1986.
- Georgopoulos, A. P., Schwartz, A. B. y Kettner, R. E. (1986). «Neuronal population coding of movement direction». *Science*, 233, 1357-1460.
- Hebb, D. O. (1949). *The Organization of Behavior*. New York, N.Y., Wiley. Ed. española, *La organización de la conducta*, Madrid, Debate, 1985.
- Johnson-Laird, P. N. (1988). *The Computer and the Mind: An Introduction to Cognitive Science*. Glasgow, William Collins Sons. Ed. española, *El ordenador y la mente*. Barcelona, Paidós, 1990.
- Jubak, J. (1992). *La máquina pensante*, Barcelona, Ediciones B. (Orig. 1991).
- Lumsden, C. J. y Wilson, E. O. (1981). *Genes, Mind and Culture*. Cambridge, Mass., Harvard University Press.
- Malsburg, C. von der y Schneider, W. (1986). «A neural cocktail-party». *Biological Cybernetics*, 54, 29-40.
- Polyshyn, Z. (1980). «Computation and Cognition: Issues in the foundation of cognitive science». *Behavioral and Brain Sciences*, 3, 111-134.
- Polyshyn, Z. (1984). *Computation and Cognition*. Cambridge, Mass., M.I.T. Press.
- Rosenblat, F. (1961). *Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms*. Washington, D. C., Spartan Books.
- Rumelhart, D. E. y McClelland, J. L. (eds.) (1986). *Parallel Distributed Processing. Cambridge. Vol I: Foundations. Vol II. Psychological and Biological Models*. Cambridge, M.I.T. Press.
- Waddington, C. H. (1957). *The strategy of genes: a discussion of aspects of theoretical biology*. London, Allen & Unwin.
- Wolfe, J. M. Cave, K. R. y Franzel, S. L. (1989). «Guided search: an alternative to the feature integration model for visual search». *Journal of Experimental Psychology*, 15, 419-433.

Aceptado el 30 de septiembre de 1993