Mentes reales

La ciencia cognitiva y la naturalización de la mente

BIBLIOTECA UNIVERSITARIA

Ciencias Sociales y Humanidades

Filosofía

Mentes reales

La ciencia cognitiva y la naturalización de la mente

Compiladores

Juan José Botero
Jaime Ramos
Alejandro Rosas

Universidad Nacional de Colombia
Facultad de Ciencias Humanas

Siglo del Hombre Editores

Mentes reales: la ciencia cognitiva y la naturalización de la mente / Jean-Michel Roy... <et al.>; compiladores Juan José Botero, Jaime Ramos y Alejandro Rosas –Santafé de Bogotá : Siglo del Hombre Editores, Universidad Nacional de Colombia, Departamento de Filosofía, Facultad de Ciencias Humanas, 2000.
244 p. ; 21 cm.

ISBN 958-665-033-2

1. Teoría del conocimiento. 2. Mente y cuerpo. I. Roy, Jean-Michel II. Botero, Juan José, comp. III. Ramos, Jaime, comp. Tít. IV. Rosas, Alejandro, comp.
121 cd 20 ed.
AGX2356

CEP-Biblioteca Luis-Ángel Arango

La presente edición, 2000

© Universidad Nacional de Colombia
Facultad de Ciencias Humanas - Departamento de Filosofía
Ciudad Universitaria Santafé de Bogotá D.C.
Tel.: 316 5000 Ext.16208

© Siglo del Hombre Editores
Cra 32 Nº 25-46 Santafé de Bogotá D.C.
Tels.: 344 0042 - 337 7738 Fax:337 7665
siglodelhombre@sky.net.co

Diseño de colección y portada
Mauricio Melo González
maurmelo@colomsat.net.co

Ilustración portada
Cabeza de Constantin Brancusi
© ADAGP París

ISBN: 958-665-033-2

Impresión
Panamericana Formas e Impresos
Calle 65 No. 94-72
Santafé de Bogotá

Impreso en Colombia-Printed in Colombia

Todos los derechos reservados. Esta publicación no puede ser reproducida ni en su todo ni en sus partes, ni registrada en o transmitida por un sistema de recuperación de información, en ninguna forma ni por ningún medio sea mecánico, fotoquímico, electrónico, magnético, electroóptico, por fotocopia o cualquier otro, sin el permiso previo por escrito de la editorial.

ÍNDICE

INTRODUCCIÓN 9

I

EL PROYECTO NATURALISTA: ENFOQUES 29
Naturalización de la mente y autonomía teórica
de la explicación mentalista 31
Jean-Michel Roy
El problema de la representación 57
Juan José Botero
Cómo evitar matar al hombre del pantano 81
David Papineau
Explicación y justificación:
hacia el naturalismo en la filosofía moral 101
Alejandro Rosas

II

EL PROYECTO NATURALISTA: OBJECIONES AL
COMPUTACIONALISMO 119
¿Es usted una máquina? 121
Garret Thomson
La intencionalidad irreductible del calcular 137
John Kearns
Simbolismo vs. conexionismo:
la estructura de las representaciones 155
Jaime Ramos

Exploraciones en ciencia cognitiva 181
Una teorizacion cognitiva acerca de las diferencias entre conceptos predicativos o relacionales y conceptos operativos o funcionales 183
Carlos E. Vasco U.
Estilos de procesamiento cognitivo 203
Christian Hederich
Neurociencias y dolor: del nociceptor a la cognición. 227
Carlos Moreno

Índice de autores 238
Los autores 241

INTRODUCCIÓN

LA CIENCIA COGNITIVA Y LA NATURALIZACIÓN DE LA MENTE

J.J. Botero/ J. Ramos/ A. Rosas

La filosofía de la mente era un terreno prácticamente abandonado hasta que el surgimiento y el desarrollo de la ciencia cognitiva permitieron retomar, con nuevos elementos, reflexiones y discusiones que durante largo tiempo habían permanecido estancadas. Puede afirmarse que esta conjunción con la ciencia cognitiva es la marca distintiva de la filosofía contemporánea de la mente .

Ahora bien: la ciencia cognitiva, en cuanto tal, intenta constituirse como ciencia natural de la mente. Su propósito puede entenderse como el de la construcción de una teoría de los fenómenos mentales que permita elaborar explicaciones que sean aceptables para un enfoque naturalista de las propiedades de la mente. Por esta razón, la mayor parte de los filósofos que trabajan en el área amplia y no delimitada con precisión denominada "filosofía de la mente" asumen una posición más o menos naturalista, en el sentido de darles a las explicaciones de los fenómenos mentales un lugar claro y coherente dentro del marco de nuestras demás explicaciones del mundo natural. Esto no quiere decir que todos busquen elaborar el mismo tipo de teorías, —pues, como veremos, puede haber discrepancias substanciales entre ellos—, sino solamente que todos buscan que sus teorías estén bien integradas en y sean pertinentes para las ciencias naturales de la cognición.

Por otro lado, la ciencia cognitiva espera que el trabajo de los filósofos aporte la claridad y la precisión conceptuales requeridas para disponer como mínimo de objetos y temas bien delimitados para trabajar. Tomemos, por ejemplo, lo que en alguna parte Fred

Dretske llama 'el problema de la soprano', y más generalmente el problema de la 'pertinencia causal' de lo mental. La soprano puede romper un vidrio con su voz mientras interpreta una canción, pero si lo hace será en virtud de las propiedades físicas de los sonidos que produce, y no en virtud de las significaciones de las palabras que entona. Éstas son 'causalmente irrelevantes', es decir, no son pertinentes para explicar lo que le ocurrió al vidrio. La pregunta es si algo semejante es cierto igualmente para los contenidos mentales en general. El comportamiento también es un evento o proceso físico, lo mismo que el rompimiento del vidrio. Si el contenido de mis creencias, deseos, y otros estados mentales, es extrínseco a la causa del comportamiento, entonces no es pertinente para su explicación causal y, por consiguiente, no debería entrar en un cuadro explicativo de lo mental que quisiera ser naturalista. Pero si sí lo es, entonces esto debería poderlo mostrar una teoría de los estados mentales desarrollada enteramente en términos naturalistas.

La claridad y la precisión conceptuales son necesarias ante todo porque los intentos concretos de 'naturalizar' el concepto de lo mental generan debates y controversias, empezando por la definición del propósito mismo: ¿Qué se debe entender por 'naturalización de la mente'? Ahora bien: aunque el consenso en la comunidad científica acerca de la necesidad de naturalizar lo mental no haya producido un acuerdo sobre cómo ha de entenderse exactamente este programa, sí es posible discernir una posición dominante al respecto. En las últimas décadas, al menos, esta posición se confunde con una u otra versión de lo que se conoce como *computacionalismo*. No es exagerado decir que el entusiasmo intelectual generado por el surgimiento, en la década de los sesenta, de lo que hoy llamamos 'ciencia cognitiva', se debió en buena parte a la creencia de que el modelo computacional proporcionaba la interpretación naturalista de la mente por excelencia. Él resolvía, finalmente, el problema, heredado desde Descartes, del lugar de la mente en el mundo natural. Este enfoque, cuya aparición más evidente es la teoría conocida como *funcionalismo*, ponía fin al dualismo, sin tener que eliminar del espectro de las ciencias a una serie de ciencias especiales, precisamente las que constituyen las llamadas ciencias cognitivas.

El modelo simbolista de la mente supone que los procesos mentales, y en particular los procesos cognitivos (percepción, aprendizaje, memoria, razonamiento, etc.), son procesos algorítmicos de manipulación de símbolos que se realizan en un medio físico complejo (el cerebro o un sistema artificial de circuitos o *chips*, como en los computadores) gracias a la 'arquitectura' funcional del sistema, es decir, a la manera como interactúan causalmente los distintos meca-

nismos del sistema. El paradigma de este modelo computacional está representado por las denominadas 'máquinas de Turing', un tipo de algoritmo diseñado por Alan Turing que es lo suficientemente poderoso para computar cualquier función computable y suficientemente simple para ser 'implementado' en una máquina (véanse los ensayos de John Kearns y Jaime Ramos, para más detalles). Desde un punto de vista lógico-matemático, las máquinas de Turing nos dieron una explicación formal y precisa de la noción intuitiva de 'algoritmo', y desde un punto de vista epistemológico y psicológico nos dieron una hipótesis plausible de cómo sistemas puramente físicos (cerebros o máquinas) son capaces de pensar y realizar conductas inteligentes. Turing nos mostró una manera de descomponer procesos inteligentes en procesos puramente mecánicos. Sin embargo, ésta no es, presumiblemente, la única manera, y desde mediados de la década pasada un modelo alternativo, el modelo conexionista o de redes neuronales, ha ido ganando cada día más seguidores. A diferencia del modelo de Turing, en el estudio de las redes neuronales, que se conoce como de procesamiento distribuido paralelo, no hay símbolos físicos atómicos, sino que la información se encuentra representada de modo distribuido en los patrones de conexión de un gran número de unidades muy simples (véanse los ensayos de Jaime Ramos y Carlos Vasco). Se ha presentado una polémica sobre si se han de entender los procesos cognitivos como simbólicos o si el modelo conexionista (inspirado en la estructura funcional del cerebro) representa un microanálisis más veraz de lo que sucede en la mente humana.

Si bien los modelos simbólicos (inspirados en los trabajos de Turing y en la inteligencia artificial, o IA) y los conexionistas (inspirados en la estructura funcional del cerebro) compiten acerca de cuál sea la mejor explicación formal de algunos procesos cognitivos importantes, ambos comparten no sólo la admisión de que existen computaciones mentales, sino también un compromiso de principio con una teoría básica acerca de la naturaleza de los estados mentales, conocida como 'materialismo'. El problema de la interacción entre lo mental y lo físico ha sido ampliamente discutido en la tradición filosófica, al menos desde Descartes, pues parece incomprensible cómo procesos puramente físicos, como los que tienen lugar en el cerebro, puedan tener efectos sobre algo supuestamente no físico como los estados mentales. Desde luego, es igualmente misterioso cómo algo puramente mental, que no tiene extensión, ni localización, ni masa, pueda tener efectos causales sobre el mundo físico. Pues bien, en el ámbito de la ciencia cognitiva es usual adoptar alguna variante de la *tesis de la identidad*, según la cual cada proceso mental es idéntico a algún proceso físico.

La versión más fuerte de la tesis de la identidad identifica cada clase de estado mental con alguna clase de estado físico. Así, por ejemplo, un sujeto se encuentra en tal clase de estado psicológico (digamos 'creer que está lloviendo') si y sólo si se encuentra en tal clase de estado físico (digamos, tener activación neuronal en tal zona de la corteza cerebral). Tal versión, conocida como fisicalismo de clases (*type physicalism*), es, sin embargo, demasiado fuerte y restrictiva porque, como persuasivamente han argumentado H. Putnam y J. Fodor, esto implicaría que todo sujeto que esté pensando que 'está lloviendo' debería tener exactamente el mismo patrón de activación neuronal exactamente en la misma zona de la corteza; pero la configuración neuronal de los sujetos es relativamente variable y es posible, como señaló Karl Lashley, que el mismo estado cerebral cumpla distintas funciones en distintos momentos, incluso en el caso de un mismo sujeto. Más aún, el fisicalismo de clases implicaría una especie de 'chauvinismo' de especie: que seres con una estructura biológica diferente de la nuestra no podrían creer que está lloviendo, porque por definición creer tal cosa es idéntico a tener activadas tales y tales de nuestras neuronas.

Como nos lo recuerda Fodor (1981, p. 127), muchos filósofos que quieren ser naturalistas abrazan por principio una postura fisicalista, expresada de modo muy general como la doctrina que postula que todos los procesos, eventos, estados de cosas existentes admiten en principio una descripción en términos físicos, pero confunden dos tesis diferentes. La primera, llamada 'de la generalidad de la física', sostiene que la física es la ciencia natural básica, y tiene el siguiente sentido: 'todos los eventos que caen bajo las leyes de alguna ciencia son eventos físicos y, por consiguiente, caen bajo las leyes de la física'. La segunda es una tesis surgida del ideal de los positivistas de la primera mitad del siglo XX, quienes buscaban la construcción de una ciencia unificada, una sola ciencia de la naturaleza en la que las distintas disciplinas (física, química, biología, psicología, etc.) estuvieran conectadas mediante leyes puente que definieran los predicados y las leyes de una disciplina superior en términos de los predicados y leyes de una disciplina más básica, siendo la física la más básica de todas. De manera confusa se piensa que esta segunda tesis es una consecuencia de la primera, lo cual le ha dado al fisicalismo el carácter de una doctrina *reduccionista* que debe regular la actividad científica. Como escribe Hempel, "todas las ramas de la ciencia tienen en principio una y la misma naturaleza y son ramas de la ciencia unitaria: la física" (Hempel, 1949). La reducción nomológica de una teoría T1 a una teoría más simple T2 tendrá la forma siguiente, según Nagel (1961): sea 1) una ley de una teoría particular T1:

1) Todos los C son (producen, causan) D

Las condiciones necesarias y suficientes para que la ley expresada por la fórmula 1) sea reducible a una ley de la teoría T2 son que las fórmulas siguientes 2) y 3) sean leyes, siendo P y Q predicados, simples o complejos, de la teoría T2:

2a) Todos los C son P
2b) Todos los D son Q
3) Todos los P son (producen, causan) Q

Y para que la totalidad de T1 sea reducible a la física se requiere que todas las leyes de T1 sean reducibles de la misma forma que 1).

Este proyecto fracasó porque los predicados de las ciencias especiales no se pueden definir en términos de los predicados de ciencias más básicas (véase Fodor [1981]).

Estos problemas con la versión reductiva llevaron a una versión distinta del materialismo conocida como materialismo de eventos (*token materialism),* que, junto con el funcionalismo, ha sido ampliamente aceptada dentro de la ciencia cognitiva. Según ella, aunque cada evento (o estado) mental es idéntico a algún evento físico, eso no significa que las clases de estados mentales se puedan reducir a las clases de estados físicos. La taxonomía de lo mental no coincide con la taxonomía de lo físico. Un mismo tipo de estado mental puede ser realizado de múltiples maneras por organismos distintos, según su propia configuración física (o incluso por máquinas, suponiendo que éstas puedan llegar a pensar, como creen los partidarios de la inteligencia artificial). Lo que determina la clase de estado psicológico *e* en que se encuentra un organismo no son las propiedades fisicoquímicas del estado cerebral, sino la función que dicho estado desempeña en la vida psíquica y conductual del individuo: las relaciones causales que tiene *e* con ciertos estímulos sensoriales (o datos de entrada del sistema, como se dice en el lenguaje computacional), las relaciones causales con otros estados mentales del sistema (creencias, deseos, etc.), y las relaciones con la conducta del sistema (datos de salida). Si, por ejemplo, distintos estados físicos llevan a uno o a distintos individuos a responder "sí" cuando se les pregunta "¿está lloviendo?", y los impulsa a buscar una sombrilla cuando han decidido salir, entonces tenemos distintas realizaciones o instanciaciones físicas de un mismo tipo de estado psicológico. Esta teoría es aplicable, en particular, a las actitudes proposicionales tales como "creer que p", "desear que p", etc. (donde p es una variable en el lugar de una proposición o sentencia). Psicológicamente, la estructura

bioquímica o física del estado funcional "cree que p" no es un dato pertinente.

Otra vertiente del proyecto naturalista adopta un *materialismo eliminacionista*, doctrina según la cual el fracaso de reducir las categorías psicológicas a las categorías biológicas es irrelevante, porque al fin y al cabo las muy antiguas categorías psicológicas (creencias, deseos, etc.) están fundadas en una mítica taxonomía del mundo, derivada de la muy ingenua psicología del sentido común que no corresponde a nada en la realidad, pues *en realidad* no hay tales cosas como deseos y creencias (véase, por ejemplo, Churchland [1979]). Podemos, por tanto, eliminar las categorías psicológicas del lenguaje científico y esperar a que una ciencia cognitiva madura se construya con base en los conceptos y el lenguaje de una neurología desarrollada. Tal vertiente radical no ha tenido una mayoritaria aceptación en el seno de la ciencia cognitiva, y muchos teóricos siguen defendiendo la necesidad y la validez de usar categorías psicológicas en la explicación de los fenómenos mentales. No todo el mundo está dispuesto a echar por la borda el legado de Freud, Vygotsky, Piaget, Chomsky y muchos otros, cuyas obras están construidas sobre los cimientos de una psicología *intencional* (véanse, por ejemplo, Fodor [1987], Dennett [1989]). El materialismo de eventos, aliado con el funcionalismo, constituye la versión dominante de la naturalización de la mente.

Pero no todos los filósofos de la mente comparten el entusiasmo por esta versión del naturalismo. Las dudas más generales dirigidas contra el *funcionalismo* afirman que un análisis funcional no es aplicable a algunas clases de estados mentales. Entre éstos se encontrarían las sensaciones como tener un dolor, o la sensación cromática al percibir una mancha roja; las emociones, como la ira; y los deseos. Pues si bien todos estos estados mentales pueden identificarse por sus causas y sus efectos típicos —sus roles funcionales— es plausible sostener que, al menos en el caso de las emociones y las sensaciones, su naturaleza está definida no sólo por las relaciones funcionales del estado con otros eventos, sino también por el aspecto cualitativo de la experiencia. De un robot de metal que pasara a un estado *n* en su programa cuando es golpeado, y dicho estado causara que dijera "¡Me duele!", y las demás conductas típicas de dolor, no diríamos que experimenta dolor o que su estado *n* sea idéntico, psicológicamente, a nuestro dolor, así sea funcionalmente idéntico. Quizá la palmaria naturaleza subjetiva del dolor haya llevado a este sector de la ciencia cognitiva a aceptar que su rol funcional no agota su explicación (sin embargo, véase el ensayo de Carlos Moreno, en este volumen, para un intento de mostrar su pertinencia cognitiva).

Si bien el caso de las sensaciones puede dejar perplejo al funcionalista, hay aproximaciones que buscan explicar las emociones y los deseos. Un entusiasta de la IA, como Daniel Dennett, argumentaría lo siguiente: si el computador puede completar unas tareas suficientemente complejas, entonces puede mostrar una conducta dirigida a metas. Eso sería razón suficiente para atribuir estados de deseo al computador. No hay ningún impedimento para hacer tales atribuciones, cuando las relaciones entre entrada y salida son suficientemente complejas. Lo mismo vale para las emociones: tienen un contenido que debe ser especificado en términos de sus roles causales o funcionales. La cuestión de si un ser o sistema puede ponerse verdaderamente enojado depende únicamente de la complejidad de las relaciones entre los estados de entrada y salida pertinentes.

Pero el problema más serio para el funcionalismo, sin duda, es si logra explicar realmente la *intencionalidad* y la naturaleza de las representaciones mentales. Quizá el padre del computacionalismo, Turing, y/o los padres del conexionismo, McCulloch y Pitts, y sus seguidores Rumelhart y McClelland, nos enseñaron cómo un proceso racional o inteligente se puede descomponer en procesos puramente físicos. Esto es, al menos, lo que quieren creer, y quieren hacernos creer, muchos filósofos entusiasmados con la IA. Ellos creen que los pequeños éxitos de la IA en la creación de sistemas físicos 'inteligentes' son un hecho científico y cultural destinado a revolucionar nuestro concepto de la mente. En estos círculos, no es una barbaridad decir que los computadores son mentes que piensan, y que los humanos somos máquinas y realizamos computaciones en el cerebro. Pero, ¿podemos *realmente* atribuir intencionalidad —pensamiento y representación— a sistema físicos, como las computadoras?

Ya el filósofo austriaco Franz Brentano, responsable, a fines del siglo XIX, de haber rescatado desde la filosofía escolástica el concepto de *intencionalidad*, sostenía que ésta es una propiedad irreducible de lo mental: sólo lo mental la tiene, y nada físico puede tenerla. Algunos autores, prominentemente Searle (1980), y en este volumen Garrett Thomson y John Kearns, piensan que es una barbaridad decir que los humanos somos máquinas o que los computadores piensan. Thompson, en "¿Es usted una máquina?", sostiene que detrás de la mente-máquina del funcionalismo computacional está el prejucio fisicalista de que lo intencional es sospechoso, porque el lenguaje intensional en el que se expresa no obedece al principio lógico de la sustitución *salva veritate*. La posición computacional busca explicar la conciencia en términos extensionales. Dado que en

los computadores la información está presente sintácticamente, podemos describir sus estados mentales de manera extensional. Cuando los describimos de modo intencional, estamos solamente usando una taquigrafía conveniente. Ésta es la posición de Daniel Dennett en *Conciencia explicada*, que Thompson escoge para ilustrar su objeción. Dennett cree que cualquier teoría científica de la mente tiene que interpretarla como un autómata, y que el modo intencional es sólo un aparato útil para entender su comportamiento. Esta posición instrumentalista se debe contrastar con una posición realista acerca de la intencionalidad, que afirma que las personas realmente creen y desean. Thomson muestra con una simple reflexión que el intento de reducir el modo intensional al modo extensional necesariamente fracasa. Aun cuando podemos hacer una representación extensional de lo intensional, hacerlo requiere del metalenguaje intensional. La información que contiene una computadora en modo extensional no es la información que una persona tiene cuando cree que p. Es entonces una confusión conceptual decir que los computadores piensan o que las mentes humanas computan.

John Kearns, en "La intencionalidad irreductible del calcular", sostiene que es un dogma falso de la lógica creer que una descripción sintáctica del lenguaje puede hacer superflua la semántica. Esto puede valer para lenguajes artificiales, mas no para los naturales. El error fundamental de los computacionalistas, que consiste en creer que un objeto o proceso mecánico puede tener intencionalidad originaria, se debe, según Kearns, a una mala interpretación de la tesis Church-Turing, según la cual toda función numérica efectivamente calculable puede ser 'computada' por un procedimiento efectivo de manipulación de caracteres. Kearns cree que hay una confusión conceptual en la base de la ciencia cognitiva, y es la suposición de que la verdad de la tesis Church-Turing implica que las máquinas pueden llevar a cabo procesos cognitivos similares a los nuestros y que, inversamente, nosotros nos comportamos mecánicamente cuando calculamos. Entre los procesos mecánicos que suceden en la máquina y los procedimientos efectivos como calcular hay similitudes, en especial si los procedimientos consisten en una manipulación de caracteres basada en los rasgos sintácticos de los mismos; es decir, en los rasgos causalmente eficaces. Pero un procedimiento, a diferencia de un proceso causal, es algo que puede ser ejecutado correcta o incorrectamente. Esto sólo lo pueden hacer las personas que ejecutan los procedimientos. Si el procedimiento se entiende como un conjunto de instrucciones que es entendido y seguido por personas, entonces estamos hablando de pensamiento. Pero si el procedimiento se entiende como lo que

sucede en la máquina, como un proceso causal, entonces es un error catalogarlo de pensamiento.

Jaime Ramos, por su parte, sostiene que la arribución de pensamiento a cerebros o máquinas se debe a una confusión categorial En "Simbolismo vs. conexionismo: la estructura de las representaciones", examina bajo esta óptica las críticas planteadas por los defensores del simbolismo al modelo conexionista, inspirado en las redes neuronales. Los defensores del modelo simbolista le atribuyen el privilegio de ser el único que puede explicar el carácter constitutivo, productivo y sistemático del pensamiento. Ramos pone en tela de juicio que estas tres características del pensamiento se den en virtud de alguna ley natural, pues es concebible que ellas aparezcan más bien en nuestra manera de hablar de él. Aun así, Smolensky ha contestado a estas críticas, sugiriendo una manera de explicar estas tres caracteristicas desde modelos conexionistas. Pero Fodor y los simbolistas no ven en las propuestas de Smolensky sino un truco para darle un carácter constitutivo a un sistema de representaciones que no lo tiene. Ramos concuerda, pero sostiene que esta misma objeción vale para los modelos simbolistas. Tampoco ellos encuentran constituyentes físicos o neurobiológicos que correspondan a los elementos constitutivos del pensamiento. Si hay razones para dudar de que los conexionistas puedan señalar procesos neurológicos que correspondan a los símbolos discretos de un modelo tensorial, las mismas razones hay para dudar de que el modelo simbólico nos proporcione una descripción veraz de tales procesos. Ramos piensa que la pregunta acerca de cuál es la verdadera estructura de las representaciones mentales se debe a la confusión conceptual de creer que el cerebro o las máquinas piensan y tienen representaciones y que, si miramos con cuidado, allí las encontraremos. La decisión de escoger entre modelos simbólicos o conexionistas para describir los procesos cognitivos se funda en razones pragmáticas, no en la supuesta correspondencia de uno u otro modelo con la forma como se realizan los procesos cerebrales.

Para otros filósofos, más dispuestos a aceptar un enfoque claramente naturalista, el debate entre simbolistas y conexionistas plantea el problema de la verdadera naturaleza y estructura de las representaciones. El estudio de la estructura y las funciones del cerebro sugiere un tipo de representación muy distinto de lo que sugerían los algoritmos de Turing. ¿Puede definirse lo que es una representación, o lo que son los procesos cognitivos, sin tener en cuenta los hechos sobre el cerebro? Jean-Michel Roy se plantea esta pregunta en su "Autonomía teórica de la explicación mentalista". Señala que la Ciencia Cognitiva está en la situación paradójica de querer naturalizar la mente y los procesos cognitivos y, sin embargo, de querer

hacerlo con autonomía teórica de la investigación del cerebro. Roy investiga la presunta autonomía teórica de lo que se denomina —por Kosslyn y Koenig— el 'enfoque seco' o computacional de la mente, propio de Fodor y de la psicología cognitiva —Simon y Newell, Neisser, Glass *et alia*— y el 'enfoque húmedo', propio de los neurocientíficos. Gracias a precisiones conceptuales en torno al concepto de autonomía teórica, Roy descubre que ya los psicólogos cognitivos admitían una dependencia lógica débil y mal definida respecto de las neurociencias. Fodor suplió posteriormente las deficiencias de esa concepción y aclaró los parámetros de la autonomía. En su concepción, la psicología cognitiva goza de una total autonomía heurística —es decir, en el orden del descubrimiento— respecto de la neurobiología. El psicólogo debe elaborar sus hipótesis sobre procesamiento cognitivo independientemente de los hechos neuronales. Pero en el momento de buscar confirmación y plausibilidad neuronal, el psicólogo debe abordar la tarea psicofisiológica de encontrar las estructuras neurales que permiten realizar los procesos de los que habla. Así, la psicología mantiene una dependencia lógica débil respecto de la neurobiología, no para su contenido, pero sí para su confirmación. Esto es lo que Fodor señaló en *Psychological Explanations* (1968). Pero este modo de abordar los procesos cognitivos ha sido cuestionado recientemente por un enfoque que se autodenomina 'húmedo'. Aunque Roy no coincide con los defensores de éste en la imputación que le hacen al enfoque 'seco' de pasar por alto completamente al cerebro, reconoce, sin embargo, que el enfoque húmedo tiene una propuesta diferente a la del enfoque seco. El enfoque húmedo limita de entrada la especulación psicológica sobre los procesos cognitivos a procesos que pueden ser realizados por el cerebro. La psicología del enfoque húmedo no se hace independiente de las estructuras cerebrales, sino que se hace teniéndolas presentes y buscando sus propias posibilidades de procesamiento cognitivo.

El problema planteado, entre otros, por John Searle en el célebre experimento del Cuarto Chino, que aparece recurrentemente en las críticas a la opción computacionalista, en particular a la versión de Jerry Fodor, es interpretado por Carlos E. Vasco, en "Una teorización cognitiva sobre las diferencias entre conceptos predicativos o relacionales y conceptos operativos o funcionales", como un dilema que se plantea, para explicar las actividades cognitivas, entre procesamiento sintáctico y comprensión semántica. Vasco intenta disolverlo mediante la aclaración y teorización cognitiva de las diferencias entre conceptos predicativos y conceptos operativos. Los primeros son, según la teoría clásica posterior a Frege y Russell, los predicados y relaciones; los segundos equivalen a las funciones de un

solo argumento, pero entendidas, no en la forma más abstracta de la matemática escolar, sino apelando a su lado 'activo', como operadores de un solo puesto (transformaciones unarias o asignaciones monádicas). Un concepto operativo corresponde a una actividad de transformación de otro concepto al cual se le aplica el primero. Vasco propone visualizar los conceptos predicativos de un puesto como 'regiones de redes semánticas' acotadas sobre una 'telaraña tridimensional' por superficies convexas que encierran una parte de ésta, 'parte en la cual quedan incluidos un número finito de nudos, nodos, núcleos o vértices contiguos, algunos de ellos conectados entre sí por cuerdas, arcos, lazos o aristas'. Esta visión holista (o más bien sisté mica) recibe el nombre de 'teoría de las regiones tridimensionales de telarañas', o 'teoría RTT'. En este dispositivo pueden ocurrir tres tipos de actividades, que explican las distinciones que ya Frege había hecho en su momento: una actividad 'propositiva', cuyo resultado es una proposición; una actividad 'judicativa', en la que se considera, evalúa y pone a prueba la veracidad de la proposición; y una actividad 'asertiva' de la proposición, cuyo producto es el acto ilocucionario de aseverar (o sentenciar) lo juzgado previamente. La diferencia entre los conceptos operativos y los conceptos predicativos tiene que ver, no con características internas de los procesos en los que intervienen para dar lugar a esas actividades, sino con lo que se va a hacer con las aserciones en las que intervienen: es decir, si las aseveraciones se utilizan como un producto que se encadena con otros, o si se las toma como productos terminales, como guías o medios de orientación y control de actividades ulteriores, propias o ajenas. En estos casos, los de los conceptos predicativos, una aserción, dice Vasco, no introduce cambios en los nodos o arcos de la red, sino que puede construir una especie de 'bandera' utilizable para la actividad concomitante a la aseveración, con propósitos de control, monitoría o elección. La actividad operativa es, pues, puramente sintáctica: su producto es otro concepto, a partir del cual se pueden iniciar otras actividades, operativas o predicativas. En cambio, el producto de la actividad predicativa es semántico, pues se sitúa, para decirlo en términos computacionales, en la interfaz entre el procesador de la información y los módulos ejecutores. La distinción entre conceptos operativos y conceptos predicativos, en consecuencia, corresponde a la distinción entre lo sintáctico y lo semántico: corresponde a la relación que se dé entre los resultados del procesamiento y los módulos que ejecutan las acciones en el sistema procesador.

Una idea semejante a la de Vasco, de relacionar la función semántica de la inteligencia con la acción, se encuentra en los desarrollos recientes de la biosemántica, o teleosemántica, en los trabajos de

Ruth Millikan y David Papineau, entre otros. En "El problema de la representación", Juan José Botero, dejándose guiar por las reflexiones de Millikan, hace unas consideraciones críticas sobre lo que, dentro de esta vertiente, se entiende por representación, una vez que se inserta la concepción funcionalista de la mente en el concepto biológico de función. La teleosemántica pretende haber resuelto, con su concepción evolucionista, una dificultad esencial que Fodor encontró en la teoría causal y naturalista de la representación. La teoría causal pretende reducir la relación semántica de representación que existe entre "A" y A a algún tipo de relación causal entre ocurrencias de objetos de clase A e instanciaciones del estado mental "A". Los trabajos de Dretske (1981) y Fodor (1987), entre otros, van en esta dirección. Fodor señala, no obstante, que la relación causal entre un evento de tipo A y un símbolo mental "A" causado por aquél, no permite determinar el contenido de las representaciones, lo que se comprueba con el hecho de que deja al contenido en un grado de indeterminación tal, que hace imposible que haya representaciones falsas. Por ejemplo, Pedro puede, en una noche oscura, tener la experiencia perceptual o el pensamiento de una rata cuando tiene un curí ante sus ojos. Si la relación causal es la relación representacional, entonces su representación de rata valdría también para curíes e incluso para la disyunción de todos los objetos que posiblemente la causen. Esta "disyuntivitis" de la teoría causal conduce, de manera contraintuitiva, a negar que existan representaciones falsas. La biosemántica sostiene que para resolver este problema es necesario especificar aun más la relación causal de la que se trata. Sostiene entonces que, para que haya representación, no basta que un elemento indique usualmente a otro, sino que es necesario además que esta indicación sea la *función* de lo que se ha de considerar como representanción. Dentro de la biología contemporánea, decir que "A" tiene como función representar a A implica decir que la copresencia de "A" con A trajo, en el pasado, beneficios al organismo y fue por ello seleccionada. ¿Pero bajo qué condiciones puede explicarse que la existencia de un evento mental traiga beneficios cuando se presenta junto con eventos a los que presuntamente representa? Botero señala, siguiendo a Millikan, que lo que determina la función de representar no es la simple conexión causal entre un evento real y un símbolo mental. Lo que determina que algo sea un símbolo es el sistema que lo usa o lo "consume". Sólo cuando el sistema que representa puede hacer algo con la información así representada, es decir, cuando puede usar la representación para producir acciones reproductivamente beneficiosas, tiene sentido hablar de representación. Lo que un evento mental representa, pues, son

las circunstancias en las que, si el evento mental está presente, se seguirán unas acciones del organismo que aumenten su éxito reproductivo. Así, por ejemplo, si tales y cuales neuronas de la corteza visual de la rana se activan ante la presencia de una mosca, entonces representan la presencia de la mosca, porque tales mecanismos perceptuales fueron seleccionados en el curso de la evolución en razón de que permitieron a las ranas iniciar acciones destinadas a obtener alimento. No obstante, al final de su ensayo, Botero plantea algunas dudas acerca de la noción misma de "función", que él considera, como Searle, un concepto relativo a un observador y, por consiguiente, que no designa ningún hecho nuevo del mundo natural. Su intervención, entonces, en la explicación será meramente instrumental, en el sentido de que no habrá nada que sea *intrínsecamente* una función biológica, por ejemplo. Y si esto es así, entonces tampoco habrá nada que sea intrínsecamente una representación. Así mismo esboza el inicio de una crítica al enfoque de la teoría de la selección natural subyacente en la teoría de Millikan, principalmente, y propone abandonar el enfoque que él llama *prescriptivo* en favor de uno *proscriptivo*: la selección opera, no prescribiendo una solución óptima, sino proscribiendo lo que es incompatible con la supervivencia y la reproducción.

David Papineau se enfrenta, en el ensayo "Cómo evitar matar al hombre del pantano", a una objeción ética contra la teoría biológica de la representación. Si, como afirma esta teoría, la inteligencia y la representación sólo pueden asentarse sobre funciones biológicas, y si éstas dependen de procesos pasados de selección natural, entonces criaturas inteligentes sin pasado biológico no pueden ser realmente inteligentes. Esto puede sonar inofensivo, hasta que uno se pregunta si es entonces lícito usarlas como carne de consumo alimentario, en el caso de que no se trate de máquinas, sino de criaturas orgánicas. Lo que está en juego son los derechos del célebre "hombre del pantano", formado por la combinación, gracias a un azar prodigioso ocurrido en un pantano, de las moléculas requeridas para conformar una entidad indistinguible de cualquier ser humano. Este problema cuenta con una salida fácil y elegante: olvidémonos de la selección natural y sus desbordados requerimientos de tiempo pasado, y adoptemos un concepto prospectivo de función, que sólo se determine por el beneficio futuro. Si las gentes del pantano se comportan realmente como criaturas inteligentes, podrán tener sus representaciones y sus derechos. Esta salida no está disponible, sin embargo, para la biosemántica, pues ella sólo admite como función aquello que tenga un pasado de selección natural. Pero Papineau encuentra una salida no menos elegante. Se trata de distinguir entre la realidad y las

meras posibilidades. Si es al menos posible que existan criaturas del pantano y que se comporten de modo inteligente, entonces habrá que aceptar que es posible un mundo donde criaturas inteligentes son creadas instantáneamente del caos por procesos especiales que no requieren de selección natural, un mundo donde el concepto prospectivo de función es aplicable. Sería sin duda un mundo muy distinto al nuestro y allí estas criaturas tendrán todos los derechos que queramos o tengamos que otorgar. Lo que no es correcto, sin embargo, es creer que estos experimentos mentales pueden cambiar en algo los resultados de la ciencia actual. Eso sería como creer que alguien puede, legítimamente, poner en duda la teoría sobre la composición química del agua apelando a la mera posibilidad de que líquidos idénticos al agua en todas sus propiedades, tengan una composición química XYZ. Lo que se necesita para hacer tambalear teorías como la biosemántica, es lo mismo que se necesitaría para cuestionar teorías químicas acerca de la composición química del agua, a saber, evidencias empíricas y no magros experimentos mentales. Tal como están las cosas, si lo que nos importa es tener una concepción adecuada de la inteligencia humana, la clave está en las funciones retrospectivas, las que se explican por procesos de selección natural.

Los ensayos de Vasco y de Papineau muestran que la inteligencia no se debe aislar de sus contextos de acción práctica si se la quiere concebir adecuadamente, y sobre todo si se la quiere naturalizar. Esto sugiere la pregunta por las consecuencias de la naturalización de la mente en nuestra filosofía moral y de la acción. Y sin embargo, precisamente en este terreno, la fuerte distinción intuitiva que acostumbramos a hacer entre explicar y justificar, entre causas y razones, parece invalidar de plano cualquier proyecto naturalista. Un proyecto explicativo naturalista sólo puede hablar de procesos causales y acaso explicar la existencia de actitudes morales, pero no podría justificarlas. Ésta es, de algún modo, la mismo objeción de Fodor contra la teoría causal de la representación: la relación de representación es normativa, pero la relación causal no lo es. Sin embargo, no parece imposible concebir un naturalismo que se proponga poner el discurso normativo en relación con el discurso causal sin ser, ipso facto, una eliminación del discurso normativo. Parte del argumento de Alejandro Rosas, en su ensayo "Hacia el naturalismo en la filosofía moral", es mostrar que no es privativo del naturalismo relacionar explicaciones causales con justificaciones normativas. La filosofía moral de Kant, por ejemplo, no se basa en la oposición de causas a razones. Opone, es cierto, causas naturales a razones, pero sólo porque adhiere a un análisis incompatibilista de la acción y coloca a la razón práctica de entrada como un fenómeno no-natural.

Pero concibe en definitiva a la razón como una causa especial, y explica a la vez que justifica la moralidad postulando la existencia de propiedades morales sui géneris, con una causalidad también sui géneris, a saber, los fines de la razón pura práctica. Así, el verdadero problema del naturalismo es si su explicación es plausible, y si funciona al mismo tiempo como justificación. No es difícil ver que la explicación naturalista ofrecida por Hume en su *Tratado de la naturaleza humana* (Lib.III, Pt.II, sec.ii), la cual describe las condiciones psicológicas y ambientales de la aparición de las actitudes morales, es una explicación que justifica y da sentido a la moralidad. Ella explica la moralidad como el medio que hace posible la cooperación entre individuos egocéntricos en un mundo de escasez moderada de recursos. Esta concepción naturalista puede incluso ser reforzada, piensa Rosas, por desarrollos recientes en la tradición de la ética evolucionista, que intentan mostrar cómo puede existir en cada individuo una actitud social y cooperativa genuinamente naturalista, que no es un artificio conscientemente inventado, sino una parte de su diseño evolutivo.

Christian Hederich, por su parte, nos presenta los resultados concretos de un experimento diseñado para sacar conclusiones sobre diferencias en el desenvolvimiento o la programación cognitiva de los estilos cognitivos clásicos. Este intento por describir las diferencias entre los estilos cognitivos clásicos en términos de la psicología del procesamiento de la información es innovador. Constituye un intento por integrar aspectos culturales y de la personalidad en el concepto de función cognitiva y, en definitiva, de ampliar las fronteras de la ciencia cognitiva hacia los contextos sociales en los que la mente está inmersa. Los estilos cognitivos denominados "independiente" y "sensible", se concibieron inicialmente en relación con el procesamiento de la información visual. El estilo sensible procesa la percepción de la vertical con una clave externa visual, mientras que el independiente usa su sentido interno de la gravedad. Posteriormente, esta distinción se generalizó a otros dominios cognitivos e incluso a dominios afectivos y sociales, y llegó a relacionársela más con la psicología de la personalidad que con la psicología cognitiva computacional. Hoy hace falta describir esto en términos de la psicología del procesamiento, y el autor decide explorar la posibilidad, sugerida por Goldberg *et alia* (1970), de entender los distintos estilos como maneras diversas como el sujeto se autoprograma para realizar una o más tareas cognitivas. Para ello diseña una prueba, denominada la prueba Logan de lectura selectiva y recuerdo, y se la aplica a 154 adolescentes de uno u otro sexo, los cuales han sido previamente seleccionados en sensibles e independientes por medio de la

aplicación de la prueba EFT (*Embedded Figures Test*). La prueba se compone de dos tareas. Inicialmente se requiere leer un relato que comprende cinco testimonios de los acontecimientos inmediatamente anteriores a un asesinato, y subrayar durante la lectura todos los nombres de muebles contenidos allí. Una vez recogidos los relatos subrayados, se requiere hacer un reporte de los acontecimientos relatados. Con base en los relatos subrayados y en los reportes de lo leído, la prueba sirve para explorar y evaluar hipótesis sobre las diferencias en las estrategias para procesar información lecto-escrita presentadas por individuos pertenecientes a estilos cognitivos distintos.

En "Neurociencia y dolor: del nociceptor a la cognición", Carlos Moreno explora otra frontera de la ciencia cognitiva. Como ya habíamos dicho, la ciencia cognitiva ha prestado escasa atención al dolor, por su presunta poca participación en procesos cognitivos. Moreno plantea que una de las posibles explicaciones de esto es la equívoca tesis neurológica según la cual la estimulación de los nociceptores es procesada a nivel subcortical, y que el dolor es percibido fundamentalmente por el tálamo, sin la participación de la corteza cerebral. De allí que se considere el dolor como un mecanismo primitivo sin mayores implicaciones para los procesos cognitivos superiores. No obstante, el autor discute evidencia reciente sobre la participación de algunas zonas corticales en el procesamiento de varios aspectos específicos del dolor, entre las que se destacan la corteza somatosensorial y la corteza cingulada. A pesar de que no se han identificado neuronas altamente especializadas en la respuesta a estímulos nocivos, el presente análisis de los hallazgos sostiene que el procesamiento cerebral de la información nociceptiva, incluyendo aspectos relacionados con la memoria y el aprendizaje, constituye un verdadero sistema funcional.

* * *

La ponencias recogidas en este volumen fueron presentadas en el Simposio Internacional de Ciencia Cognitiva realizado en Bogotá en el mes de marzo de 1998. Fue organizado por el Departamento de Filosofía de la Universidad Nacional de Colombia, el cual agradece aquí a las instituciones que brindaron su apoyo: Colciencias, la misma universidad a través de la Facultad de Ciencias Humanas y la vicerrectoría académica, y la Hemeroteca Nacional Universitaria. Su publicación es posible gracias al auspicio de la División de Investigación de la Sede de Bogotá y de la Facultad de Ciencias Humanas. Se espera con ella hacer una contribución filosófica a ese campo interdisciplinario, relativamente reciente, en el que participan desde

neurofisiólogos y psicólogos hasta lingüistas y filósofos, y que se conoce como ciencia cognitiva. Los editores del volumen hemos organizado los ensayos presentados aquí en tres secciones, teniendo en cuenta dos criterios principales. En primer lugar, ubicamos en una misma sección los ensayos que comparten, así sea de una manera laxa, una misma orientación filosófica general con respecto al proyecto naturalista de la ciencia cognitiva. En segundo lugar, hemos ordenado los ensayos de tal manera que vayan primero aquellos que expliquen nociones y argumentos que ayuden a entender los subsiguientes. En la primera sección, titulada "El proyecto naturalista: enfoques", se incluyen los trabajos de Jean-Michel Roy, Juan José Botero, David Papineau y Alejandro Rosas, los cuales abordan diversos aspectos del problema de expresar en términos de ciencias naturales explicaciones intencionales y normativas. En una segunda sección, titulada "El proyecto naturalista: objeciones al computacionalismo", hemos incluido los trabajos de John Kearns, Garrett Thompson, y Jaime Ramos, que contienen críticas a la interpretación de los procesos mentales en términos de computaciones efectuadas sobre representaciones mentales codificadas neuralmente. Y en una tercera sección, titulada "Exploraciones en ciencia cognitiva", se incluyen los trabajos de Carlos Eduardo Vasco, Christian Hederich y Carlos Moreno, quienes presentan propuestas para abordar procesos cognitivos específicos en el marco general de una ciencia natural de la mente.

BIBLIOGRAFÍA

CHURCHLAND, P. M. (1979), *Scientific Realism and the Plasticity of Mind*, Cambridge: Cambridge University Press.

DENNETT, D. (1989), *The Intentional Stance*, Cambridge (Mass.): MIT Press. Trad. de Daniel Zadunaisky (1991), *La actitud intencional*, Barcelona: Gedisa.

DENNETT, D. (1991), *Consciousness Explained*, Boston: Little, Brown & Co.

DRETSKE, F. (1981), *Knowledge and the Flow of Information*, Cambridge (Mass.): MIT Press.

FODOR, J. (1981), "Special sciences", reimp. en *Representations*, Cambridge (Mass.): MIT Press. Trad. de Jesús Fernández Zulaica (versión ampliada) (1985), "Dos clases de reduccionismo" en: J. Fodor, *El lenguaje del pensamiento*, Madrid: Alianza Editorial.

HEMPEL (1949), "The Logical Analysis of Psychology", en H. Feigl and W. Sellars (eds.) *Readings in Philosophical Analysis*, New York: Appleton Century Crofts.

NAGEL, E. (1961), *The Structure of Science*, New York: Harcourt, Brace and World. Trad. de Néstor Miguez (1981), *La estructura de la ciencia*, Barcelona: Paidós.

SEARLE, J. (1980), "Minds, Brains and Programs", *The Behavioral and Brain Sciences*, vol. 3. Trad. de Guillermina Feher de la Torre (1994), "Mentes, cerebros y programas", en: M. Boden, *Filosofía de la inteligencia artificial*, México: Fondo de Cultura Económica.

SMART, J.J. C. (1959), "Sensations and Brain Processes", *Philosophical Review*, vol. 68.