

# PSICOLOGIA ORDINARIA y CIENCIAS COGNITIVAS

Según una concepción muy amplia, son “cognitivas” todas las disciplinas que se ocupan del modo en que los sistemas, naturales o artificiales, registran, almacenan y procesan una determinada información y actúan en el mundo en función de ella.

La “revolución cognitiva” en psicología ha tomado a su cargo la rehabilitación de conceptos de los que desconfiaba el conductismo, como los de “representación mental”, “creencia”, “conocimiento”, “intención” y tantos otros.

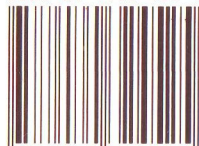
La psicología cognitiva, y también la IA, otorgarían legitimidad científica a los conceptos que utilizamos corrientemente para explicar el comportamiento de nuestros semejantes y el de otras criaturas, cuando les atribuimos la posesión de diversos estados mentales.

En esta obra se consideran los problemas epistemológicos y ontológicos que plantean la elaboración y transformación de los conceptos psicológicos corrientes en el marco de las ciencias cognitivas. ¿Cuál es el estatuto de las “representaciones”, “conocimientos”, etc. utilizados en este tipo de explicación? ¿Cuál es la naturaleza de los estados “mentales” cognitivos?

En este libro se reflejan con claridad las tensiones producidas en el interior de este nuevo campo de investigación en el intento de dar respuestas a tan acuciantes problemas metateóricos.

Todos los textos provienen de ponencias presentadas en un seminario sobre “Filosofía del espíritu” coordinado por Pascal Engel en el marco de un programa sobre “Ciencias de la Comunicación” del CNRS en Francia.

ISBN 84-7432-452-1



9 788474 324525

gedisa  
editorial

Código 2.388

Colección Hombre y Sociedad

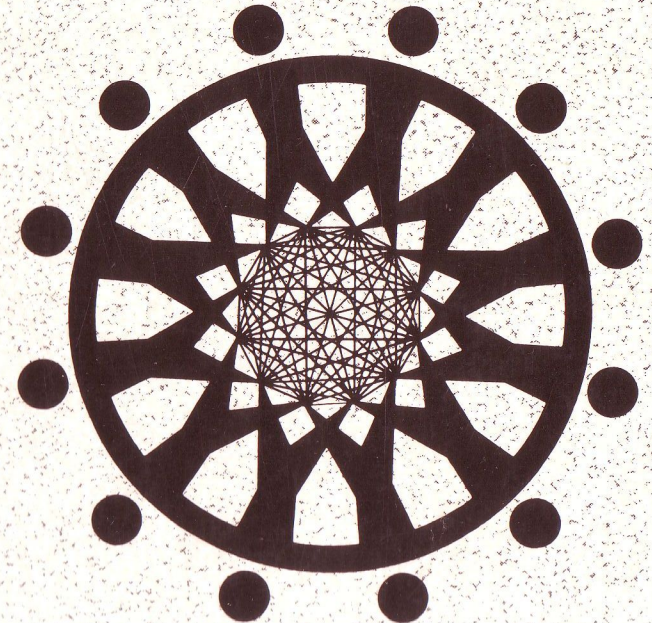
Serie

CIA·DE·MA

P.V.P. Bs.  
2100

F156

# Pascal Engel (Comp.) PSICOLOGIA ORDINARIA y CIENCIAS COGNITIVAS



gedisa  
editorial

De la colección de PAPELES JPG

Para http://padron.entretemas.com/vea ORDINARIA y CIENCIAS COGNITIVAS

Pascal Engel (comp.)

---

**PSICOLOGIA ORDINARIA  
Y CIENCIAS COGNITIVAS**

De la colección de **PAPELES JPG**  
en <http://padron.entretemas.com.ve>

**Grupo: PSICOLOGÍA**  
**Subgrupo: CIENCIAS COGNITIVAS**

Editorial Gedisa ofrece  
los siguientes títulos sobre

## CIENCIAS COGNITIVAS

pertenecientes a sus diferentes  
colecciones y series  
(Grupo "Psicología")

PASCAL ENGEL (COMP.) *Psicología ordinaria y ciencias  
cognitivas*

PAUL M. CHURCHLAND *Materia y conciencia.  
Introducción contemporánea  
a la filosofía de la mente*

F. VARELA, E. THOMPSON  
y E. ROSCH *De cuerpo presente*

DANIEL C. DENNETT *La libertad de acción*

HILARY PUTNAM *Representación y realidad*

DANIEL C. DENNETT *La actitud intencional*

FRANCISCO J. VARELA *Conocer*

JEROME BRUNER *Realidad mental y  
mundos posibles*

# PSICOLOGIA ORDINARIA Y CIENCIAS COGNITIVAS

por

**Pascal Engel (comp.)**

De la colección de **PAPELES JPG**  
en <http://padron.entretemas.com.ve>

**gedisa**  
editorial

Título del original francés: *Psychologie ordinaire et sciences cognitives*  
© Centre National de la Recherche Scientifique, París, 1988

Traducción: Alberto Luis Bixio  
Diseño de cubierta: Gustavo Macri

Primera edición, Barcelona, España, 1993

Derechos reservados para todas las ediciones en castellano

© Editorial Gedisa  
Muntaner, 460, entlo., 1ª  
tel. 201 60 00  
08006 - Barcelona, España

ISBN: 84-7432-452-1  
Depósito legal: B - 5819/1993

Impreso en Libergraf  
Constitució, 19

Impreso en España  
Printed in Spain

Queda prohibida la reproducción total o parcial por cualquier medio de impresión, en forma idéntica, extractada o modificada, en castellano o cualquier otro idioma.

## Indice

1. ¿Puede la psicología cognitiva apelar a la psicología ordinaria?, por <i>Pascal Engel</i> .....	9
2. Variedades de la representación mental, por <i>Andrew Woodfield</i> .....	28
3. Psicología del sentido común y ciencias cognitivas, por <i>Samuel Guttenplan</i> .....	49
4. Actitudes mentales y psicología del sentido común, por <i>Radu J. Bogdan</i> .....	72
5. El conocimiento tácito, por <i>Martin Davies</i> .....	112
6. ¿Tienen efectos físicos los fenómenos mentales?, por <i>Daniel Laurier</i> .....	144
7. Un dilema para el fisicismo, por <i>Howard Robinson</i> .....	170
GLOSARIO.....	190
DEBATES.....	194
Cognitivistas, todavía un esfuerzo, por <i>Mario Borillo</i> .....	195
Observaciones sobre las ciencias cognitivas, por <i>Jean Pierre Desclès</i> .....	200
El informe sobre las ciencias cognitivas, por <i>Michel Imbert</i> .....	205

Las dos dimensiones de la cognición, por <i>Jean François Richard</i> .....	207
¿Qué es lo cognitivo?, por <i>Jean-Paul Codol</i> .....	212
¿Qué es lo cognitivo?, por <i>Guy Tiberghien</i> .....	221

De la colección de **PAPELES JPG**  
en <http://padron.entretemas.com.ve>

# 1

## ¿Puede la psicología cognitiva apelar a la psicología ordinaria?

*Pascal Engel*

“Cuando hablamos de la ‘psicología como ciencia natural’ no debemos suponer que esta expresión designa una clase de psicología fundada definitivamente en un suelo firme. Quiere decir exactamente lo inverso; quiere decir una psicología particularmente frágil y en la cual las aguas de la crítica metafísica se escapan por todas las juntas, una psicología cuyas hipótesis básicas y datos deben reconsiderarse partiendo de conexiones más amplias y traducidas en otros términos. Estas, en suma, la expresión de una falta de confianza, no expresión de arrogancia... Unos cuantos hechos crudos, un poco de cháchara y de especulación sobre opiniones, un poco de clasificación y de generalización en el nivel solamente descriptivo, el prejuicio que *tenemos* de los estados de espíritu y de que nuestro cerebro los condiciona... eso no es ciencia, es sólo la esperanza de una ciencia. La mejor manera en que podamos facilitar ‘una realización científica’ es comprender hasta qué punto es grande la oscuridad en que nos movemos y no olvidar nunca que las hipótesis de las ciencias naturales de las cuales hemos partido son hipótesis provisionales y sujetas a revisión.”

William James, *Principles of Psychology*, 1982, pág. 467 (citado por Pylyshyn, 1984, pág. 272)

Los artículos reunidos aquí tratan todos, en un aspecto u otro, sobre las relaciones entre la psicología ordinaria y la psicología cognitiva. En esta presentación mi finalidad es tratar de explicar por qué el problema planteado por la naturaleza de esas relaciones (problema que recientemente ha ocupado el centro de una vasta bibliografía) es a la vez un problema metodológico central para lo que se ha dado en llamar “ciencias cognitivas” y para la filosofía del espíritu y de la psicología en general.

Cuando en la vida de todos los días explicamos la conducta de nuestros semejantes, utilizamos toda una variedad de térmi-

nos psicológicos tales como “creer”, “dudar”, “desear”, “temer”, etc. Naturalmente suponemos que esos términos denotan estados mentales internos de los sujetos a quienes se los atribuimos y que esos estados tienen cierto contenido que expresamos habitualmente con frases como “Juan cree que Fidel es un sabio” o “Juan desea que Khadafi se calme”. Según la terminología usual de los filósofos se puede llamar a esos estratos mentales *actitudes proposicionales*, porque suponen una actitud del sujeto en relación con un contenido proposicional (que *Fidel es un sabio, que Khadafi se calme*). Suponemos también que esos estados o actitudes proposicionales son de una manera u otra la *causa* de la conducta y de las acciones que observamos: Juan se fue a vivir a Cuba *porque* creía que Fidel es un sabio; Juan dirigió un ultimátum a Khadafi *porque* deseaba que Khadafi se calmase. La forma más corriente de estas explicaciones causales ordinarias implica en el agente la presencia conjunta de creencias (o por lo menos de estados en virtud de los cuales el agente tiene cierta información sobre el mundo o supone que existe un determinado estado de cosas) y de deseos. Creencias y deseos interactúan para causar las acciones: Juan fue a Cuba porque creía que Fidel es un sabio y porque deseaba ir a vivir a Cuba. Esos estados tienen asimismo un poder de predicción: dada esa creencia y dado ese deseo, predecimos que el agente irá a vivir a Cuba. Desde luego, suponemos también que muchas creencias o deseos pueden interactuar y que son causas de una conducta únicamente si otras creencias o deseos no ponen obstáculos a aquellos que atribuimos al agente: si por ejemplo el agente cree también que la vida en Cuba carece de comodidades y si esa creencia es suficientemente fuerte para impedirle desear ir a vivir allí, no haremos la predicción de que la creencia de que Fidel es sabio y el deseo de ir a Cuba son suficientes para llevar al agente a vivir en Cuba. Este esquema común de explicaciones es en gran medida irreflexivo y espontáneo: forma parte de la concepción intuitiva que tenemos de nosotros mismos, de nuestras acciones y de las acciones de nuestros semejantes. De conformidad con la terminología ahora usual entre los filósofos, se puede suponer que este esquema está en el centro de la psicología ordinaria o psicología del sentido común (*folk psychology*). Entonces se plantean dos cuestiones importantes:

1) ¿ En qué medida se puede suponer que la psicología

ordinaria es una teoría o un esquema unificado de explicaciones y de principios y cómo describir ese esquema?

2) ¿ Puede una psicología científica evolucionada invocar legítimamente el mismo tipo de explicaciones y principios que las explicaciones y principios que, según se supone, forman la psicología ordinaria?

Supongamos por el momento que la psicología ordinaria sea ciertamente ese esquema unificado. Hasta fines de la década de 1950, la psicología científica del siglo XX estuvo dominada por el conductismo. Según los conductistas la psicología ordinaria es falsa: no hay estados *mentales* denotados por los términos habituales de actitudes proposicionales, porque esos estados son disposiciones (físicas) tendientes a una conducta observable. Los hombres no creen ni desean sino que obran porque han estado sometidos a cierta intensificación de las respuestas producidas por la presencia de repetidos estímulos y no porque haya fenómenos mentales internos. Una psicología científica conductista es pues incompatible con la psicología ordinaria. Pero, la mayor parte de los psicólogos conductistas no se preocupaba por saber si sus investigaciones tenían o no alguna relación con la psicología ordinaria; se contentaban con ignorarla y escribían sus artículos sobre el trabajo que realizaban en los laboratorios con su vocabulario conductista. Skinner (1971) es una excepción: sostiene Skinner que el conductismo muestra simplemente que términos tales como “libre”, “digno”, “responsable” no corresponden a los seres humanos y que, por lo tanto, nuestra concepción de nosotros mismos es radicalmente falsa. De manera que frente a la psicología ordinaria Skinner adopta la posición que habrá de llamarse *eliminatória*: sencillamente debemos desembarazarnos de nuestro vocabulario psicológico mentalista habitual. La posición adoptada por gran número de filósofos (como Ryle) en la década de 1950 es diferente: según ellos, los términos mentalistas usuales pueden *redefinirse* con la *misma significación* (o se puede mostrar que en realidad tienen la misma significación) que los términos que designan aquellas disposiciones a una determinada conducta. Por ejemplo, “tiene miedo” tiene el sentido de “estar uno dispuesto a adoptar una conducta de fuga”. Este “conductismo lógico” es una doctrina que se refiere a la *significación* de los términos mentales de la psicología ordinaria y dicha doctrina sostiene que, una vez hechas las traducciones “conductistas”, la

psicología ordinaria torna a ser compatible con una psicología científica conductista. Desgraciadamente esta doctrina es falsa. Y es falsa no sólo porque hay numerosos estados mentales (como las sensaciones) que no parecen poder definirse por alguna conducta observable específica, sino también porque aun cuando sea posible una definición en términos conductistas, ésta sólo puede formularse mediante la postulación de *otros* estados mentales, los cuales a su vez no pueden definirse desde el punto de vista conductista sino en virtud de otros estados mentales, y así sucesivamente (véase por ejemplo, Davidson 1980).

En el mismo momento en que los filósofos cobraban conciencia del fracaso del conductismo lógico, la psicología científica se apartaba del conductismo en general. Una gran parte del comportamiento, especialmente el humano, no puede analizarse desde el punto de vista conductista. Lo que en psicología se ha dado en llamar la “revolución cognitiva”, a partir de la década de 1960 (pero que ya había comenzado mucho antes), consistió en volver a introducir en el seno de las teorías experimentales de la psicología los conceptos mismos que los conductistas consideraban irremisiblemente sospechosos porque eran “mentalistas”: por ejemplo, los conceptos de *creencia*, de *conocimiento*, de *plan*, de *fin*, de *deseo*. Basta con abrir una obra de psicología cognitiva (por ejemplo Neisser 1967 o Johnson - Laird - Wason 1977) para comprobarlo. La noción de “cognición” es evidentemente difícil de definir y todos aquellos que apelan a ella no quieren decir necesariamente lo mismo cuando la emplean. Pero uno de los sentidos fundamentales de este concepto es el de que los organismos entran en contacto con su medio y muestran ciertas conductas porque son capaces de adquirir ciertas informaciones o ciertos conocimientos sobre ese ambiente, es decir que son capaces de representárselo de diversas maneras. Uno de los postulados básicos de la psicología cognitiva es pues, el de que existen ciertos estados internos de los organismos que se pueden tratar como *representaciones mentales*, las cuales permiten a esos organismos almacenar cierta *información* sobre el ambiente; dicha información, combinada con los planes o los fines de tales organismos, es la causa de su conducta. Como veremos, todo depende de la manera en que se definan nociones tales como información, representación mental o fin. Pero es llamativa la analogía con la manera en que

la psicología ordinaria postula estados mentales provistos de contenido como causas de la conducta.

La analogía es tan llamativa que numerosos filósofos y psicólogos cognitivos pudieron considerar que en psicología y en ciencias cognitivas las explicaciones no eran sustancialmente diferentes de las explicaciones que da la psicología del sentido común en la vida de todos los días cuando se trata de dar cuenta de las acciones de los agentes. Según esos filósofos y psicólogos, la psicología cognitiva puede valerse del mismo tipo de explicaciones intencionales del comportamiento humano que la psicología ordinaria y explican ésta mediante actitudes proposicionales que poseen las siguientes propiedades:

I) Son semánticamente evaluables, es decir, tienen contenidos intencionales que poseen condiciones de verdad;

II) están dotadas de poder causal sobre otras actitudes y sobre la conducta;

III) son de tal condición que las generalizaciones implícitas de la psicología ordinaria tocantes a creencias y deseos son en gran medida verdaderas.

Fodor (1975, 1981, 1985, 1987) y Pylyshyn (1984, 1987) son los principales representantes de esta tesis. Según ellos, una explicación en psicología cognitiva (y también las otras ciencias cognitivas) debe, lo mismo que las explicaciones de la psicología ordinaria, suponer que los agentes obran en virtud de estados intencionales y de representaciones dotadas de contenidos que son las causas de sus actos. Las formas usuales de esas explicaciones dicen, por ejemplo, que una criatura hace *X* porque ha inferido *P* sobre la base de la *creencia q* y hace *X* porque dicha criatura había realizado esa inferencia y tenido esa creencia y porque deseaba (proyectaba, quería, etc. ) que *X* fuera hecho. Para que haya explicación es menester que haya por lo menos una ley o una regularidad legal. Según Fodor y Pylyshyn, porque hay generalizaciones cognitivas que tienen esta forma: Si *X* cree en *p* y desea *q*, entonces, *mutatis mutandis*, *X* hace *A*, puede explicarse esa conducta de tipo *A*. Pylyshyn (1984) sostiene que cuando puede darse una explicación de este tipo en el caso de un sistema que trata de información, ese tratamiento es “cognitivamente penetrable”. La penetrabilidad cognitiva no es otra cosa que el criterio de la explicación cognitiva, según esta concepción. Porque las generalizaciones cognitivas comparten esta forma básica con las generalizaciones del sentido común se

puede decir, según Fodor, que la psicología cognitiva puede invocar el mismo tipo de explicaciones intencionales que el sentido común. Sin embargo esto representa sólo la mitad de la cuestión. La otra parte no tiene que ver mucho con el sentido común. Cuando éste hace esas generalizaciones no atiende en modo alguno a la índole de las representaciones implicadas por los estados intencionales, que el propio sentido común postula, y supone tan sólo que existe una relación causal *cualquiera* entre los estados mentales y entre los estados mentales y el comportamiento. En cambio, la psicología cognitiva procura especificar qué tipos de representaciones están presentes y cuál es su poder causal. Según Fodor y Pylyshyn, las actitudes sólo pueden satisfacer las condiciones I, II, III mencionadas si: a) se postula que los organismos tienen ciertas actitudes que son relaciones con representaciones mentales y que esas relaciones son relaciones *funcionales* o *computacionales*, y si b) los procesos mentales que están en la base de esas representaciones son series causales de inscripción de esas representaciones en el espíritu. a) corresponde a la tesis *funcionalista* de la filosofía del espíritu, según la cual los estados mentales tienen papeles causales -funcionales en el interior de un sistema de entradas y salidas comparables a los estados internos de un ordenador que realiza un programa (Fodor 1975-1987, Putnam 1975). Ahora bien, los ordenadores pueden también suministrar el modelo de la manera en que “estados” internos de la máquina pueden entrar en relaciones causales, pues las operaciones de la máquina consisten por entero en transformaciones de símbolos.

Al efectuar esas operaciones, la máquina es sensible solamente a las propiedades sintácticas (o formales o estructurales) de los símbolos.

Y las operaciones de la máquina sobre los símbolos se limitan por entero a alteraciones de sus formas (Fodor 1987, pág. 19).

El ordenador, postula Fodor, es pues la metáfora o el modelo apropiado para concebir la naturaleza del espíritu, de sus procesos, de sus representaciones y de sus poderes causales. Según esta *teoría representacional de lo mental* (TRM), las representaciones son codificadas, almacenadas y manipuladas en la forma de símbolos internos o de un lenguaje comparable con el “lenguaje mecánico” de los ordenadores y están provistas de

poderes causales *en virtud de sus formas y de su sintaxis*. Fodor postula que ese lenguaje debe ser innato, privado, propio de cada tipo de criatura y que es un “lenguaje del pensamiento” (Fodor 1975, 1987). Dejaré aquí de lado esta última tesis, cuyo sentido intuitivo bastará para nuestros fines. Fodor y Pylyshyn sostienen que esas hipótesis están de conformidad en gran medida con toda la corriente “clásica” de las investigaciones realizadas en ciencias cognitivas y que esa teoría *representacional* o *computacional* del espíritu es la teoría que está en la base de lo que podemos llamar el “cognitivismo clásico”. Pylyshyn defiende, por ejemplo, la tesis de que hay un medio único de representaciones, compuesto de símbolos proposicionales (de frases de un lenguaje interno), que es suficiente para explicar todas las representaciones cognitivas, incluso aquellas que a primera vista no parecen ser proposicionales, como las representaciones perceptivas y las imágenes mentales (Pylyshyn 1984; véase Engel de futura aparición).

Por supuesto, no se puede atribuir al sentido común hipótesis tan elaboradas y refinadas. Cuando los representantes del cognitivismo clásico pretenden invocar generalizaciones explicativas “ordinarias” no hablan de esta parte refinada de la “teoría” común, sino que se refieren al *esquema* general creencia/deseo que según ellos está en la base de la teoría. De manera que sobre este punto hay ya una ambigüedad en la tesis de que la psicología cognitiva “conserva” el sentido común. Ha de entenderse que la psicología cognitiva conserva la forma pero no necesariamente los contenidos específicos de sus explicaciones, es decir, las hipótesis I, II, III. En efecto, aun dentro de este marco clásico hay toda clase de problemas que hacen dudar de que el sentido común psicológico pueda volver a encontrar en dicho marco sus “inspiraciones” intuitivas. En primer lugar, como lo muestra A. Woodfield en este mismo volumen (véase el cap. 2), el sentido común simpatiza a priori con la idea de que hay no sólo una sino varias formas de representaciones mentales: conceptos, percepciones, imágenes, representaciones verbales o proposicionales, cada una de las cuales opera en un nivel distinto y con formas distintas. La tesis de Fodor y Pylyshyn según la cual un solo medio de representación —proposicional— subyace en todos los otros no está, pues, de acuerdo con el sentido común.



Woodfield muestra además que las recientes investigaciones realizadas en psicología de la percepción tienden antes bien a confirmar este punto de vista del sentido común. Esto ya muestra, si se acepta este parecer, que hay muchas maneras de invocar el sentido común o que la psicología cognitiva “clásica” del estilo de Fodor y Pylyshyn no está necesariamente confirmada en todas sus hipótesis básicas. En segundo lugar, una de las nociones centrales de la psicología cognitiva es la de que la cognición procede de un nivel representacional no consciente, “infrapersonal” o “subdoxástico” en virtud del cual las informaciones son almacenadas y tratadas. Las actitudes proposicionales invocadas en este nivel no son las actitudes ordinarias de creencia, deseo, etc. Son actitudes “tácitas” o “implícitas” y el tipo de “conocimiento” que ellas consideran se ha llamado clásicamente (sobre todo después de los trabajos de Chomsky y de sus discípulos), “conocimiento tácito o implícito”. Martin Davies examina aquí (véase el cap. 5) la índole de esos conocimientos y muestra que no pueden colocarse en el mismo plano que los conocimientos proposicionales ordinarios, especialmente el conocimiento reflexivo que los hablantes tienen de las reglas de su lenguaje. Sin embargo esto no amenaza, en mi opinión, la tesis de Fodor y Pylyshyn, pues esa tesis, según vimos, no implica que las generalizaciones cognitivas sobre el conocimiento tácito de las reglas lingüísticas, por ejemplo, sean las mismas generalizaciones del sentido común. Pueden ser “cognitivamente penetrables” y conservar así y todo la forma I, II, III, aun cuando el sentido común no las reconozca. Desde el punto de vista que nos ocupa, más problemático es el hecho de que Davies recurra al concepto fodoriano de *modularidad*. Es esta una noción compleja, controvertida (véase Garfield 1987), pero que se puede caracterizar en general del modo siguiente: de conformidad con la concepción arquitectónica del espíritu de Fodor (1983), hay que distinguir dos clases de sistemas autónomos en la cognición: los sistemas “periféricos” (*input systems*) como los sistemas relacionados con una modalidad sensorial (visión, audición, etc.) y el sistema “central”, que corresponde a los procesos de pensamiento y de control. Los primeros sistemas se caracterizan por un tratamiento aislado (“encapsulado”) de la información y son en este sentido *módulos*, en tanto que el sistema central se caracteriza por una integración de las informaciones, por un “holismo” de las inferencias y las representa-

ciones. Davies sugiere que la modularidad podría ser la verdadera esencia subyacente de los estados de conocimiento tácito. Fodor (1983) sugiere que el dominio por excelencia de la cognición es el de los sistemas modulares de tratamiento de la información y enuncia una “ley” (1983, pág. 140): cuanto más global es un proceso cognitivo (es decir, cuanto más central y menos “modular” es) menos se sabe a su respecto. Ahora bien, por definición, los sistemas centrales son los que producen actitudes proposicionales en el sentido habitual, tales como creencias, deseos, pensamientos, etc. Debería seguirse de ello que el dominio *propio* (o en todo caso el dominio más susceptible de conocerse) de la cognición es el de los sistemas periféricos modulares y no el de las actitudes proposicionales. Como el esquema explicativo de la psicología del sentido común apela exclusivamente a estos últimos y no a los primeros, parecería que la psicología cognitiva no fuera un buen modelo de una ciencia de la cognición. También aquí esta dificultad puede resolverse si se distinguen dos sentidos en el concepto de “actitud proposicional”: el sentido de estados conscientes del sistema central y el sentido de los estados simbólicos inconscientes (o por lo menos aquellos a los que se tiene acceso) que entran en las diversas relaciones computacionales postuladas por la teoría representacional de lo mental y la concepción computacional de Fodor. Si se hace esta distinción, las “actitudes” de que habla Fodor (cuando éste dice que la psicología cognitiva satisface las condiciones I, II, III) no son las mismas de que habla el sentido común. En rigor de verdad y en el seno de las ciencias cognitivas, únicamente la teoría de la decisión —si ella forma parte de esas ciencias—, que habla de creencias, de deseos y de preferencias más o menos conscientes de los agentes, podría apoyarse en la psicología de las actitudes del sentido común. Esto muestra (como observa S. Guttenplan) que el esquema de la psicología ordinaria a que se remite Fodor está en alto grado idealizado.

Aunque la posición metateórica de Fodor y Pylyshyn haya sido lo suficientemente influyente para que se la considerara con frecuencia como la concepción establecida, ha sido puesta en tela de juicio. En primer lugar, se la cuestionó desde el punto de vista metafísico en nombre del dualismo. H. Robinson (cap. 7) muestra que la construcción “funcionalista” de los estados mentales no logra dar cuenta adecuadamente de las sensacio-

nes, de las “cualidades segundas” ni de los estados subjetivos. En la medida en que las nociones de esos estados fenomenológicos forman buena parte del aparato conceptual de la psicología del sentido común, se puede afirmar que el cognitivismo clásico las deja de lado y que aquí afronta una dificultad muy seria. Luego aquella posición fue atacada por psicólogos propiamente dichos; ya he mencionado los argumentos aportados por Woodfield contra la idea de un sistema único de representaciones mentales y lo cierto es que muchos psicólogos admiten tales argumentos (en el caso de las imágenes, Kosslyn 1980; en el caso de las representaciones lógicas Johnson-Laird 1983). Pero la forma más radical de impugnación procede de ciertos filósofos que pusieron en tela de juicio el conjunto de los postulados I, II, III, en los cuales reposa la teoría y además el postulado general de que la psicología del sentido común suministre la forma general de las explicaciones cognitivas. Esos filósofos defienden todos en diversos grados lo que se ha dado en llamar la tesis “eliminadora”, de inspiración materialista, según la cual debería quedar lisa y llanamente *eliminado* el vocabulario mental e intencional de la psicología ordinaria en lugar de ser explicado por una teoría representacional y computacional de la cognición o reducido a los términos de ésta. Siguiendo la clasificación adoptada por R. Bodgan se pueden distinguir por lo menos tres formas de esta postura eliminadora.

La más radical es la de P. Churchland (1981, 1984, 1988) y de P. S. Churchland (1986). Según estos autores, la psicología ordinaria es, por cierto, de conformidad con lo que dice Fodor, una teoría empírica, realista y descriptiva de lo mental. Pero como tal es falsa. Es falsa porque la explicación cognitiva debe poder reducirse en última instancia a una explicación neurofisiológica en la cual la idea misma de estado mental y de contenido de un estado mental no tiene sentido. Tarde o temprano las neurociencias reemplazarán a la psicología y todo el vocabulario mental que ahora ella utiliza será en definitiva reemplazado por historias de neuronas, de sinapsis etc. que explicarán por completo lo que llamamos erróneamente representaciones mentales.

Stich (1983) defiende otra versión de la postura eliminadora. Stich está dispuesto a admitir la concepción funcionalista de lo mental propia de Fodor y también la individuación sintáctica de los estados mentales entendidos como inscripcio-

nes de símbolos en el espíritu. Pero rechaza la tesis I del carácter semánticamente evaluable de las creencias, de los deseos y de otras actitudes corrientes. Al estudiar las maneras en que el sentido común atribuye contenidos intencionales, Stich muestra que semejantes contenidos son en el mejor de los casos vagos, indeterminados e incapaces de recibir condiciones claras de verdad. En este sentido no hay una correspondencia sistemática entre las explicaciones ordinarias de la conducta y las explicaciones cognitivas. Como se ve, Stich no hace en cierto modo más que radicalizar las observaciones antes expuestas sobre las diferencias que hay entre los dos tipos de explicaciones. Según él, esos desfases son suficientemente importantes como para que no sea lícito afirmar que la psicología cognitiva pudiera apelar al sentido común.

Dennett (1978, 1987) defiende la versión más débil de la postura eliminadora (por lo demás, sostiene que no se trata verdaderamente de postura eliminadora; véase 1987 *e infra*). No admite que la psicología ordinaria sea una teoría de lo mental falsa porque no es una teoría o, más exactamente, no es una teoría empírica. Se trata de un esquema normativo idealizado (por lo tanto ni verdadero ni falso), que rige nuestras atribuciones usuales de estados intencionales cuando interpretamos la conducta de toda criatura. Igual que cualquier forma de interpretación de la conducta, la psicología cognitiva puede valerse de ese esquema pero con la condición de saber que éste no es una “instancia intencional”, heurística e instrumental que permita explicar las acciones (las de un ser humano, las de una máquina o las de un animal). Pero no cabe afirmar que los estados mentales postulados por el esquema son reales: hasta es probable, conjetura Dennett, que una psicología evolucionada no llegue a encontrar en los individuos estados correspondientes a lo que corrientemente llamamos creencias, deseos o anhelos. En el mejor de los casos una explicación cognitiva podrá reducir esos estados a conjunciones de estados infrapersonales (que se pueden presumir tácitos) que ya nada tienen que ver con las creencias y actitudes ordinarias. Sin embargo éstas conservan el poder de *racionalización* que tienen en toda interpretación.

Otra forma de impugnación de la tesis fodoriana nació dentro del campo filosófico y es la que procede de los partidarios de una concepción *externalista* de la intencionalidad. Esta forma no se ha tratado en los artículos siguientes de este libro, pero

yo diré brevemente algo sobre ella. Fodor, que en esto está de acuerdo con numerosos psicólogos, sostiene que las explicaciones cognitivas deben obedecer a lo que él mismo llama el “principio del solipsismo metodológico”, principio que establece que los únicos contenidos a los que debe apelar una psicología científica son contenidos internos e individuales de los organismos sin hacer referencia a los estados del ambiente exterior al sujeto (Fodor 1980). El problema está en que las habituales atribuciones de contenidos de actitudes proposicionales no parecen acomodarse siempre a este principio “individualista” o “cartesiano”: muy frecuentemente nuestras atribuciones apelan a entidades del mundo exterior a los sujetos y no se puede evaluar su contenido sino suponiendo que semánticamente hacen referencia a esas entidades. Eso ocurre con numerosas atribuciones que implican el empleo de expresiones ostensivas, como por ejemplo “Juan cree que este líquido es agua”. Aquí no podemos reducir el pronombre demostrativo “este” a una descripción de propiedades del agua que sean internas de la psicología individual de Juan. Putnam (1975) ha hecho sobre este tema una serie de experiencias de pensamiento tendientes a mostrar que en las atribuciones usuales de actitudes había una ambigüedad sistemática entre la atribución de estados mentales internos del sujeto (de conformidad con el solipsismo metodológico) y las atribuciones de estados semánticamente evaluables por la especificación de condiciones de verdad externas. Burge (1979, 1986) ha generalizado y extendido esta tesis a todos los tipos de pensamiento al sostener que todos los contenidos intencionales sin excepción dependen del ambiente en cuanto a su individuación. Esto amenaza la concepción fodoriana de la explicación cognitiva no sólo porque la semántica de los contenidos intencionales no sería, en esta concepción “externalista”, dependiente de su sintaxis, sino además porque la psicología cognitiva, si ésta es aún posible, debería hacer sistemáticamente referencia a las propiedades del ambiente. El sentido común quedaría a “salvo”, pero no en la forma querida por Fodor.

Como se ve, todos estos debates sobre la relación o falta de relación supuesta entre psicología ordinaria y ciencias cognitivas dependen en gran medida de las respectivas concepciones que se tengan de la psicología ordinaria y de las ciencias cognitivas.

Ya es hora pues de retornar a la primera pregunta formulada antes: 1) ¿Es una teoría la psicología ordinaria, como lo suponen tanto los partidarios de la teoría representacional del espíritu como sus adversarios, partidarios de la eliminación? ¿Hay un conjunto de leyes (aun *ceteris paribus*) más o menos permanentes que gobiernen nuestra concepción intuitiva y preteórica de lo mental y que entren en las explicaciones habituales de la conducta? Podemos dudar por las razones mismas que invocan los propios partidarios de la postura eliminatoria: si existen semejantes leyes, éstas deben poder formularse en términos de relaciones causales entre contenidos (como “si X tiene sed y si cree que esto es agua, X beberá esto”); pero si los contenidos intencionales son indeterminados, como lo sostiene Stich, ¿cómo se podrán formular esas leyes? Desde este punto de vista, es contradictorio admitir que hay una teoría presente si ésta no es especificable. Si la psicología ordinaria es una teoría, cabe dudar de que ella sea tan unificada, coherente e integrada como las teorías de otros dominios (las teorías físicas, biológicas). Como observa Guttenplan, si es una teoría, lo es en un sentido muy laxo que autoriza numerosos cambios en los conceptos básicos, y no una teoría permanente, “establecida”, como pretende Churchland. En este sentido, ¿cómo se puede decir que la psicología de los griegos era la misma que la nuestra? ¿Qué habría dicho Eurípides si se le hubiera explicado que su Medea era víctima del estrés? Esto muestra por lo menos que, si existe una psicología ordinaria o corriente, ella evoluciona. Por ejemplo, nosotros que vivimos en el siglo XX tal vez estemos más dispuestos a explicar las conductas atendiendo a “situaciones” y a “contextos histórico-psicológicos” que cierta psicología ordinaria del pasado, que se inclinaba más fácilmente a considerar rasgos de carácter o propiedades naturales y ahistóricas de los individuos. Si en su mayoría, los seres humanos actuales son capaces de comprender las reacciones de los personajes de un dibujo animado de Disney o de Tex Avery, ¿hasta qué punto pueden comprender los estados mentales de personajes de Balzac o de Stendhal, quienes, por lo que sabemos, no hablan de marcianos? ¿Y no resulta ridículo —si no es dentro del contexto de una arrogancia científica que algunos consideran típicamente norteamericana— afirmar que nuestra psicología cognitiva evolucionada podría explicar mejor la conducta que la psicología de Sófocles? Lo menos que se puede decir

es que a ciertos filósofos norteamericanos no les vendría mal seguir un curso de historia de las mentalidades. Esto no quiere decir que ese concepto de “mentalidad” no plantee problemas metodológicos importantes en este tipo de trabajos; pero es seguro que la ingenuidad preteórica no está donde generalmente se cree. Puede sin embargo dudarse de que la psicología ordinaria sea una teoría por otras razones. Como lo subraya Bogdan, lo que se llama psicología ordinaria es ante todo un esquema práctico de explicaciones y no exclusivamente una teoría y acaso no sea en modo alguno una teoría. Tal vez (como sugirió A. Morton, 1980) la psicología ordinaria no es más que una colección de recetas (ajustadas permanentemente a las sucesivas necesidades del trato con nuestros semejantes en la vida corriente) “individualista”, en el sentido de que dicha psicología no se basa ni en leyes ni en conceptos generales. Tiene uno el sentimiento de que, aun cuando se cumpla la predicción de P.M. y P.S. Churchland (de que el vocabulario *teórico* de la psicología ordinaria pueda “desaparecer” en el futuro en beneficio del vocabulario de las neurociencias), esto no significaría la desaparición de la capacidad práctica de explicar y predecir las conductas, capacidad en la que parece apoyarse ese vocabulario. La postura eliminatoria incurriría en este sentido en un verdadero error de categoría.

Podríamos extender indefinidamente estas observaciones, lo cual muestra que el concepto de “psicología ordinaria” es, como dice Guttenplan, un rótulo extraordinariamente ambiguo sobre el cual es muy difícil basar un argumento. Si existe algo por el estilo en una forma unificada es ante todo la tarea de la psicología social, de la antropología cultural e histórica y quizá de una disciplina que todavía no está presente y que sería la antropología cognitiva (véase Sperber 1985). Mientras no dispongamos de una concepción coherente de lo que pudiera ser la psicología del sentido común, no vemos bien cómo se podría valerse de ella en la explicación cognitiva o cómo se podría rechazarla; de suerte que todo el debate parece viciado.

Hay sin embargo una manera de reducir este debate dentro de límites razonables. Para hacerlo (si se admite que la psicología ordinaria es algo así como una teoría) hay que distinguir, como hace Guttenplan, entre, por una parte, las teorías y conceptos psicológicos particulares en virtud de los cuales en esta o en aquella época, en estas o en aquellas

circunstancias un individuo o un grupo puede verse llevado a explicar comportamientos (por ejemplo, la teoría y los conceptos psicoanalíticos en sus formas vulgares de inconsciente, fantasía, complejo, etc., teoría que está más o menos integrada en nuestras concepciones corrientes actuales) y, por otra parte, el núcleo teórico de las usuales explicaciones de la conducta. A mi juicio, Fodor tiene razón al hacer hincapié en que ese núcleo reposa en la atribución de estados mentales intencionales a los individuos en la forma de generalizaciones del tipo de aquellas que apelan a creencias, deseos y otras actitudes. Por ejemplo, sería por completo extraordinario comprobar que una cultura dada no atribuya estados mentales y en lugar de éstos enuncie “explicaciones” conductistas como “Juan tenía una disposición intensamente reforzada a la conducta de mover sus piernas ante estímulos X” o materialistas como “sus neuronas XYZ fulguraron” para describir la huida de alguien ante un gran perro o para manifestar un tierno pensamiento por la amada. Es posible que esta segunda manera de hablar se base, como lo sostienen los partidarios de la postura eliminatoria, en un mito (aun cuando admitan que se trate de un mito útil), así como Ryle (1949) denunciaba en nombre de una forma de conductismo lógico, el mito del “fantasma en la máquina”. Pero es muy dudoso que se puedan encontrar culturas que hubieran llevado a cabo semejante género de reformas. La cuestión no es tanto la de saber si el núcleo intencional de la psicología ordinaria —es decir, recurrir a las actitudes para explicar la conducta— es verdadero o falso como la cuestión de saber si puede uno prescindir de él. Desde este punto de vista, según lo hemos ya señalado, la psicología de los psicólogos es mixta: a veces recoge nociones que son moneda corriente, como los conceptos de creencia, de plan o de deseo y a veces forja y acuña sus propios conceptos como el de sistema modular o de transductor. Cuando dicha psicología recoge las nociones habituales les cambia por lo menos el sentido. Sin duda esto ocurre con toda disciplina científica, lo cual no significa ni una adhesión a la psicología ordinaria ni un rechazo de ella, como lo dice muy bien Pylyshyn (1984, pág. XX):

“Si es cierto que la psicología ordinaria es una mezcla de generalizaciones, luego una teoría cognitiva desarrollada contendrá sin duda cierta cantidad de térmi-

nos de la psicología ordinaria, pero dejará de lado buena parte de ellos y agregará sus propias construcciones. No veo cómo el desarrollo de esta imagen de conjunto pueda anticiparse en detalle, salvo acaso en psicolingüística y en visión. No veo, por lo tanto, por qué se deba emprender una defensa de la psicología ordinaria en general ni por qué deba negársela (véase Stich 1983)."

Esta conclusión moderada será mi conclusión oficial sobre el debate y supongo que habrá de ser en general la conclusión de los psicólogos. Pero, como filósofo, quisiera también proponer una imagen apenas más especulativa.

Las posiciones de Churchland y de Dennett antes recordadas muestran que no estamos obligados a aceptar la tesis de Fodor respecto de la causalidad de los estados mentales ni respecto de la realidad de las actitudes y de los contenidos intencionales. Churchland, como materialista y de conformidad con la concepción habitual de la causalidad, sostiene que únicamente sucesos físicos pueden ser causas. Fodor, como se recordará, resuelve el problema suponiendo que los estados mentales son relaciones con símbolos *físicos* inscritos en el lenguaje de un organismo. Pero hay otra forma posible de materialismo, diferente de la de Churchland y de la de Fodor. Es la defendida por Davidson (1980) y analizada en este libro por D. Laurier (cap. 6). Davidson sostiene que todo hecho mental concreto es idéntico a un hecho físico neurológico concreto, pero niega que haya leyes estrictas (deterministas) que relacionen *tipos* o propiedades de sucesos mentales con tipos o propiedades de sucesos físicos. Esta doble tesis (la tesis de la identidad de los *tokens* físicos y mentales) y la de la armonía de lo mental (la imposibilidad de que existan leyes psicofísicas como leyes de lo mental) es lo que Davidson llama *monismo anomal*. Se ha sostenido que esta tesis implicaba una forma de epifenomenalismo, es decir, de ineficacia de lo mental. D. Laurier muestra que no es así y que los sucesos mentales continúan siendo causas dentro de esta concepción (de acuerdo con la tesis de Davidson 1980). Esta posición es importante porque muestra que uno no está obligado a admitir una forma fuerte, reduccionista, de materialismo, como Churchland. Esta tesis en apariencia es compatible con la postura eliminatória

porque los únicos hechos que existen son físicos. Pero no se trata de una postura eliminatória sin más. En efecto, Davidson rechaza la idea fodoriana de que las generalizaciones descriptivas de la psicología ordinaria sean legales ("nomológicas"). Pero no por eso las considera estériles, como Churchland. En otras palabras, del hecho de que los sucesos mentales son sucesos físicos (particulares) no se puede llegar a la conclusión de que la psicología intencional ordinaria sea reductible a las neurociencias, así como no se puede llegar a la conclusión de que existan leyes de lo mental como tal. Según Davidson, el esquema o el núcleo de las explicaciones intencionales de la conducta es indispensable en toda interpretación de la conducta. Ese esquema no es de índole descriptiva, sino que es de naturaleza normativa, en el sentido de que toda interpretación y explicación psicológica que se dé en las "ciencias del comportamiento" o en las "ciencias morales", para hablar como antaño, debe presuponer un grado mínimo de racionalidad en las criaturas que uno interpreta (véase Engel 1989). Ese grado mínimo está expresado por las atribuciones de deseos y de creencias y por el carácter causal de tales deseos y creencias. Sobre este particular, Davidson coincide con la posición de Dennett: la psicología ordinaria tiene la índole de una norma de racionalidad. No se puede prescindir de ella no sólo porque es "útil" incluso en una teorización psicológica más evolucionada, sino también porque su esquema de conceptos es una condición a priori de toda interpretación. Esto no implica que la psicología ordinaria deba ser "eliminada" o "conservada". Implica tan sólo que los estados y los contenidos descritos por la psicología ordinaria no son reales en el sentido en que Fodor sostiene que lo son. Ahora bien, y según vimos, resulta muy difícil (contrariamente a lo que parece sostener Fodor) tomar al pie de la letra los estados intencionales de la psicología como descriptivos de los procesos y representaciones analizados por la psicología cognitiva. Sin duda tenemos necesidad de una teoría más precisa sobre la manera en que atribuimos corrientemente actitudes proposicionales a nuestros semejantes. No es seguro que basten solamente las consideraciones normativas (Gordon 1986), pero parece dudoso que podamos interpretar la conducta de los demás sin que intervengan a lo menos ciertas normas de racionalidad, aunque éstas no sean tan fuertes como Davidson y Dennett puedan decirlo. Por eso la concepción aquí esbozada

parece la más capaz de resolver las tensiones y de conciliar algunas de las intuiciones contradictorias que tenemos en cuanto al estatuto de una ciencia de lo mental (que todavía no existe) y de una psicología ordinaria (que tal vez no exista).

### Referencias bibliográficas

- Burge, T. 1979 "Individualism and the mental", *Midwest Studies in philosophy, IV*.
- Churchland, P.M. 1981 "Eliminative materialism and propositional attitudes", *Journal of Philosophy*.
- 1984 *Matter and consciousness*, MIT Press. [Hay versión castellana: *Materia y conciencia*. Barcelona, Gedisa, 1992.]
- 1988 "Folk psychology and cognitive science", *Proceedings of the Aristotelian Society*, Supp. Vol.
- Churchland, P.S. 1986 *Neurophilosophy*, MIT Press
- Davidson 1980 "Mental events" (1969), en *Essays on actions and events*, Oxford U. Press
- Dennett 1978 *Brainstorms*, MIT Press
- 1987 *The intentional stance*, MIT Press. [Hay versión castellana: *La actitud intencional*. Barcelona, Gedisa, 1991.]
- Engel 1986 "L'anomalie du mental", *Critique*
- 1988 "Questions vives en philosophie de l'esprit", *Préfaces*, 10 (en prensa) *Images mentales et psychologie cognitive*
- 1989 *La norme du vrai*, Gallimard
- Fodor 1975 *The language of thought*, MIT Press. [Hay versión castellana: *El lenguaje del pensamiento*. Madrid, Alianza editorial, 1985.]
- 1980 *Representations*, MIT Press
- 1983 *The modularity of mind*, MIT Press, tr. fr. Minuit 1985. [Hay versión castellana: *La modularidad de la mente*. Madrid, Morata, 1986.]
- 1985 "Fodor's guide to mental representation; the intelligent Auntie vade mecum"
- 1987 *Psychosemantics*, MIT Press
- Garfield 1987(comp.) *Modularity in knowledge representation and language understanding*, MIT Press.
- Gordon 1986 "Folk psychology as simulation", *Mind and Language*, 1, 2
- Horgan, T. y Woodward, J. 1985 "Folk psychology is here to stay", *Philosophical Review*
- Johnson-Laird, P, 1983 *Mental models*, Cambridge University Press
- Johnson-Laird, P. y Wason, P. (comps.) 1977, *Thinking*, Readings in cognitive science, Cambridge university Press.
- Kosslyn, S. 1980 *Image and mind*, Harvard V. Press.

- Laurier, D. 1987 "L'individuation des états intentionnels et le statut de la psychologie populaire", *Recherches sur la psychologie et le langage*, 8, Grenoble.
- Morton, A, 1980 *Frames of mind*, Oxford university Press
- Neisser, U. 1967 *Cognitive Psychology*, Nueva York
- Putnam, H. 1975 "The meaning of meaning", *Philosophical Papers*, vol. 2, Cambridge university Press
- Pylyshyn, Z. 1984 *Computation and cognition*, MIT Press
- Ryles, G. 1949 *The concept of mind*, Hutchinson, Londres, tr. fr. Payot 1976. [Hay versión castellana: *El concepto de lo mental*.]
- Sperber, D. 1985 "Anthropology and psychology: towards an epidemiology of representations", *Man* 20
- Stich, S, 1983 *From folk psychology to cognitive science* MIT Press

## Variedades de la representación mental

*Andrew Woodfield\**

La finalidad de este artículo es defender una tesis empírica: que las percepciones y los conceptos reposan en sistemas distintos de representaciones mentales. Se trata de un objetivo limitado; no diré gran cosa de las complejidades de la relación que hay entre contenido estrecho y contenido amplio. Supondré, sin dar argumentos, que las representaciones mentales, lo mismo que las otras clases de representaciones, tienen dos dimensiones: un aspecto sintáctico, que tiene que ver con su forma y sus relaciones mutuas, y un aspecto semántico, que tiene que ver con la referencia de las representaciones y con sus condiciones de verdad. Los datos empíricos utilizados aquí provienen de fuentes diversas. Si hubiéramos de ser estrictos y escrupulosos en el plano del método habría que separar dos niveles de indagación. Una cosa es preguntar: ¿Hay buenas razones, fundadas en el sentido común, para postular diferentes sistemas de representación? Pero es otra cosa muy distinta preguntar: ¿Cuáles son las razones científicas para hacerlo más allá de los datos de la introspección y del sentido común? Como filósofo, no tengo autoridad para pronunciarme sobre la segunda pregunta. Procuraré atenerme a la primera. Pero los recientes progresos registrados en la ciencia de la percepción son bien conocidos, de manera que me resulta difícil mantener los dos niveles separados. Quizás eso no tenga demasiada importancia en este caso porque la respuesta a las dos preguntas es afirmativa. Recientes trabajos realizados en ciencias cognitivas fortalecen la hipótesis de que existen numerosos sistemas de re-

\* En la edición original francesa, este artículo fue traducido del inglés por Pascal Engel [E.]

presentación y, a mi juicio, esa hipótesis es asimismo compatible con la psicología ordinaria.

El sentido común nos dice que la función de la percepción es presentar al entendimiento una información continuamente actualizada sobre el ambiente presente. La facultad perceptiva comprende diversas modalidades de percepción, cada una de las cuales está vinculada con los órganos sensoriales que le corresponden y cuya actividad depende de los estímulos llegados del exterior. La experiencia perceptual es una respuesta consciente y original a esos estímulos. En condiciones de seguridad y de calma el pensamiento puede desplegarse libremente con total independencia de lo que ocurre en las proximidades del sujeto pensante.

Los contenidos de la percepción son subjetivamente diferentes de los pensamientos. Según la útil expresión de Peacocke (1983), las experiencias perceptuales tienen propiedades de sensación ("sensacionales"), además de sus propiedades "representacionales". Los pensamientos no poseen propiedades sensoriales. Cada modalidad tiene su propio registro distintivo de propiedades sensoriales: Se puede decir, partiendo de las cualidades que percibe, si su experiencia es visual, auditiva, gustativa, etc.

Cada modalidad se limita al registro de las propiedades del mundo exterior que ella puede representar. Una experiencia visual puede registrar que un objeto tiene un determinado color, pero no que emite un sonido; la experiencia auditiva puede revelar un ruido pero no que haya algo que tenga color. Ciertas propiedades pueden representarse por obra de varias modalidades, como por ejemplo la forma, la textura o el tamaño. Una experiencia táctil de una forma dada tiene propiedades de sensación completamente diferentes de las propiedades que tiene una experiencia visual de la misma forma. Resulta igualmente muy significativo en este contexto el hecho de que el tacto y la vista puedan dar impresiones que están en conflicto; por ejemplo, un objeto puede parecer a la vista liso y sentirse al tacto como rugoso.

Fortalece nuestra tesis el hecho de que haya un registro limitado de propiedades que son representables en la experiencia perceptual pura, en tanto que el registro de lo que es representable en el pensamiento es más amplio. Si la capacidad

representacional de la percepción es más limitada que la capacidad de producir conceptos, ello se debe tal vez a que emplean diferentes sistemas de representación. El hecho de que cada modalidad se limite a su propio registro distintivo de propiedades (aun cuando haya intersecciones) sugiere también que cada modalidad tiene su propio sistema de representación, su propio repertorio y sus propias facultades. Y el hecho de que un objeto pueda parecer liso cuando se lo siente al tacto como rugoso sugiere que los mecanismos de la visión y del tacto funcionan de manera autónoma. Esos mecanismos no cotejan sus informaciones antes de decidir qué representación habrán de seleccionar en sus respectivos almacenes.

Este argumento se ve aun más reforzado si se consideran ciertas ilusiones perceptivas. Ocurre a veces que los sentidos nos engañan y que un sujeto consciente de la ilusión obre cognitivamente con miras a remediarla. En ilusiones geométricas como la de Muller-Lyer, una flecha parece más larga que la otra, por más que las dos tengan la misma longitud. Y la ilusión persiste después de haber descubierto S la verdad. Para explicar esta persistencia cabe formular la hipótesis de que la experiencia visual de S y la creencia de S reciben la mediación de representaciones distintas de las figuras. ¿Por qué? Consideremos el caso en que X tiene la experiencia visual de que un objeto O es F y simultáneamente juzga que O no es F. S no es irracional, si bien sabe que los dos contenidos coexisten en su conciencia pues puede poner entre paréntesis el contenido perceptual y puede abstenerse de considerarlo verdadero. En cambio, un sujeto que cree que O es F y simultáneamente cree que O no es F, se ve obligado a abandonar una u otra de estas creencias. Y el estado correspondiente al hecho de juzgar que O no es F parece apenas posible psicológicamente si O se presenta en el mismo modo de representación en los dos juicios. La mejor descripción del estado mental correspondiente, es decir, “saber que uno está teniendo la experiencia de una ilusión”, consiste en decir que dos representaciones que pertenecen a diferentes sistemas están activadas y que esos sistemas tienen cada uno sus funciones propias. Las dos representaciones están presentes en el espíritu de S porque cada una de ellas es el mejor producto a que puede llegar su mecanismo propio de conformidad con sus principios de funcionamiento propio.

En las situaciones perceptuales en que no se manifiesta ninguna ilusión, si S cree que O es F, descubre la anomalía y realiza una verificación para saber cuál de las representaciones es la correcta. Los dos sistemas tienen conexiones funcionales entre sí; el juicio final de S corresponde a principios de coherencia independientes de los sistemas. Sin embargo, en situaciones de las que se sabe que son ilusiones, los principios son revocables. La apariencia refractaria no desaparecerá por sí sola y S no tiene la posibilidad de hacerla desaparecer.

Cada uno de los cinco sentidos es capaz de producir una experiencia del mundo exterior que no coincide con lo que S sabe de ese mundo. Este hecho, junto con el mencionado antes de que las dos modalidades pueden producir contenidos que entran en conflicto entre sí, refuerza la tesis según la cual cada modalidad tiene su propio sistema de representaciones y de que todas ellas son diferentes del sistema conceptual.

Recordemos tan sólo algunos otros hechos bien conocidos y concentrémonos en el ejemplo de la visión. Una experiencia visual presenta a S una escena organizada espacialmente y la apariencia de la escena existe en el espacio visual. S ve no solamente objetos que tienen un color y una forma sino que ve también los espacios vacíos que quedan entre esos objetos. Si la experiencia es rica, S tendrá varios objetos representados que ocupan una variedad de relaciones espaciales enlazadas entre sí. Pero en el pensamiento los objetos no aparecen como están localizados en el espacio, pues la representación conceptual no dispone de las propiedades “sensacionales” requeridas para presentar semejante apariencia. La apariencia visual es topográfica; la representación conceptual no lo es.

Una experiencia visual momentánea puede presentar un haz de colores dispuestos en matices finamente distintos y no todos ellos tienen un nombre preciso en el vocabulario de S. En principio, en el caso de todo matiz de color que pueda distinguirse perceptivamente, S habría podido tener un concepto que correspondiera exactamente a ese matiz. En la práctica, S probablemente dispondrá de muchos menos conceptos que el número de matices que es capaz de distinguir regularmente. Observaciones semejantes se aplican a la distinción de las texturas, de las formas, de los sonidos, de los gustos, etc. Además, dos percepciones de la misma escena en diferentes condiciones de iluminación presentarán los objetos ya con un

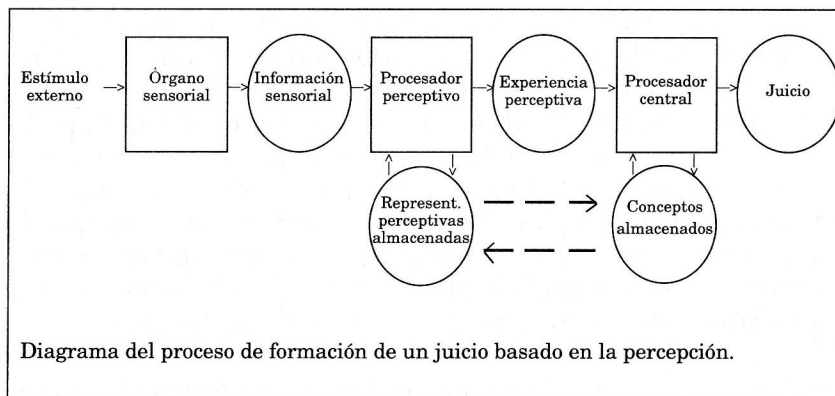


aspecto, ya con otro, y estas diferencias en los contenidos son aquellas de que se tiene experiencia. De manera que la individuación de las percepciones por los contenidos es extremadamente fina; hay muchas más distinciones en la experiencia que en el pensamiento y las descripciones de aquello que tiene que ver con la experiencia son siempre lacónicas.

Otro fenómeno significativo es el de la “memoria icónica” (Neisser 1967; Sperling, 1960), es decir, la breve retención de una impresión visual. Una vez desaparecida la escena, durante un breve tiempo el sujeto puede dirigir su atención sucesivamente a diferentes partes de la imagen memorizada y puede ver propiedades de la escena como si continuara contemplándola. Este proceso de barrido de las imágenes (*scanning*) lleva su tiempo. El sujeto observa, conceptualiza y finalmente describe aspectos de la escena que ya estaban representados visualmente desde el principio. Esos aspectos estaban todos simultáneamente en la presentación visual inicial.

La experiencia perceptual es rica y específica. El pensamiento es pobre y abstracto.

Estos rasgos introspectivos de la experiencia sensorial, así como muchos otros, fueron abundantemente estudiados por autores pertenecientes a la tradición fenomenológica, como Husserl y Merleau-Ponty, y por psicólogos experimentales. En consecuencia la psicología ordinaria sanciona implícitamente cierta imagen de esa parte del espíritu que puede representarse en un diagrama de proceso (*flow chart*), si bien la expresión “diagrama de proceso” pertenece a la jerga semitécnica; véase el diagrama siguiente.



En el diagrama, las casillas rectangulares representan mecanismos funcionalmente definidos y los círculos, tipos de representaciones. Las flechas no punteadas representan el tratamiento o procesamiento en marcha, el flujo de los procesos causales que se desarrollan en la producción de un juicio perceptual. (Por razones de simplicidad, aquí sólo se muestra una modalidad sensorial. Desde cierto punto de vista el diagrama está demasiado determinado. Muestra tan sólo un sistema de representaciones en el caso de una modalidad. En verdad, la psicología ordinaria deja abierta la posibilidad de que haya varios sistemas en el seno de una modalidad dada, cada uno de los cuales está dedicado a una parte diferente de la tarea.) No hay una flecha que indique retroacción partiendo del círculo de los conceptos utilizados en el momento del juicio y que vaya hacia el círculo que contiene la representación perceptual activada en ese momento. El sistema perceptual selecciona una representación de su almacén sin esperar la luz verde proveniente de los pensamientos de ese momento de S. Este puede quedar sorprendido; no siempre ve lo que espera ver y puede ser víctima de ilusiones. Esto representa una parte de lo que Fodor (1983) quiere decir cuando afirma que los sistemas periféricos son “modulares”: están “informacionalmente encapsulados”.

Las flechas punteadas que pasan entre los dos “almacenes” indican la posibilidad de conexiones que no están comprendidas en el tratamiento de la información en marcha. Cada almacén contiene el repertorio de representaciones latentes de S. La manera en que S puede ver cosas o pensar cosas está limitada al tiempo  $t$  por los dos almacenes en  $t$ . El registro de las posibilidades ofrecidas a S tiene sus raíces, en definitiva, en las necesidades biológicas; S debe tener suficientes conceptos de especies de objetos que le permitan representarse estados del mundo de manera tal que S pueda sobrevivir. En el curso de la vida, los contenidos de los almacenes pueden cambiar, pueden adquirirse nuevas facultades de representación, ya sea por maduración, ya sea por aprendizaje y los dos almacenes influyen el uno en el otro. Evidentemente el proceso de armonización es de mediano o de largo plazo si el repertorio perceptual disponible en  $t$  fue influido en parte por experiencias y pensamientos anteriores; en cierto sentido las percepciones están limitadas de manera muy general por expectativas. Esto último

no va contra la tesis de la modularidad, siempre que las creencias no interfieran en el tratamiento de la información en marcha.

En el momento del acto de percibir, el procesador perceptual busca en el almacén representaciones perceptuales y selecciona una (o sintetiza un nuevo elemento entre los que están almacenados) que sea acorde con la información sensorial que entra. Cuando la representación seleccionada es activada da una experiencia. Ella puede también hacer que el procesador central seleccione dentro del almacén de conceptos uno o varios conceptos que son utilizados en un juicio consciente. Estos procesos de selección y de activación están indicados por flechas verticales en el diagrama.

Si la imagen esbozada aquí es en general correcta, ciertas doctrinas filosóficas están erradas. Algunos filósofos, por razones epistemológicas o metafísicas, decidieron negar o minimizar la diferencia que hay entre la percepción y el juicio. Han sostenido, por ejemplo, que todo acto de percibir es un juicio perceptual (Armstrong 1963) o han abrazado una versión exagerada de la idea de que la percepción está siempre cargada de teoría. Esta tendencia puede verse todavía en P. Churchland (1979), cuando dice: "La percepción consiste en la explotación conceptual de la información natural contenida en nuestras sensaciones o en nuestros estados sensoriales." (pág. 7) <sup>1</sup>

Muchos filósofos han caído en la trampa de una conceptualización exagerada de la percepción al poner acento excesivo en el papel de las *descripciones* o reseñas de la percepción de un sujeto. Esas descripciones, por el hecho de ser verbales, están en realidad conceptualizadas; son expresiones de juicios basados en la percepción. Strawson (1979), por ejemplo, se propone sostener que las percepciones no son datos o razones para juzgar que una escena es esta o aquella. Trátase de una observación válida: desde el punto de vista de S, no hay inferencia del hecho de ver al hecho de juzgar en los casos corrientes, pero Strawson trata de establecerlo al sostener que las experiencias de S están cargadas teóricamente desde el principio. La razón que aduce Strawson sencillamente no es válida. <sup>2</sup>

En esto los psicólogos cognitivos parecen tener más conciencia de la dificultad. Por ejemplo, Foster (1985) observa (y Fodor en 1985 está de acuerdo con él) que "el acceso a las informaciones de salida del sistema periférico (*input system*) se realiza a través de la UTC (Central Processing Unit, Unidad de

Procesamiento Central)". Strawson no considera la posibilidad de que la experiencia sensorial pueda ser capaz de presentar estados objetivos (y no ya siempre los datos sensoriales) y ser ello no obstante no conceptual.

Hintikka (1975) sostiene, contra una interpretación de Husserl, que las percepciones son percepciones de rasgos objetivos y que sin embargo "ninguna *noesis* separada se requiere para la individuación perceptual...; una percepción inédita es tan intencional y está tan articulada como, por ejemplo, un pensamiento". Pero yo no estoy de acuerdo con Hintikka cuando éste dice que las percepciones están conceptualmente articuladas de manera espontánea: "La conceptualización se elabora directamente partiendo de esas mismas impresiones sensoriales no reconstruidas" (*ibid.*, págs. 202-203). No me parece que Hintikka tenga un argumento que le permita descartar la idea de percepciones puras, que también están articuladas pero no conceptualmente. Su argumento, como el de Strawson, me parece enteramente basado en el hecho indudable (que sin embargo aquí no es pertinente) de que las *descripciones* de percepciones implican una conceptualización.

John Searle (1983) sostiene que "a veces cierto predominio conceptual es una condición previa de la experiencia visual" (pág. 54).

La figura  $\overline{1001}$  puede "verse como la palabra TOOT, como una mesa con dos grandes pelotas debajo, como el número 1001 con una línea trazada encima, como un puente con dos cañerías debajo, como los ojos de un hombre que lleva un sombrero con un bramante que pende de cada lado y así sucesivamente. En cada caso tenemos una experiencia diferente por más que sean constantes los estímulos visuales, las líneas trazadas en el papel y la luz que ellas emiten" (*ibid.*). Son todas experiencias diferentes, convengo en ello, pero ninguna de ellas es pura. Hay sin embargo una presentación de la forma, del tamaño y de los colores de la imagen impresa en la página y esa presentación es pura. Pero no presupone un predominio conceptual. Un sujeto que tiene aptitudes conceptuales puede interpretar la figura, y las diversas interpretaciones posibles tienen sus raíces en la misma experiencia pura. Normalmente los sujetos no hacen una inferencia consciente que vaya de la percepción pura a la interpretación, pero existe un nivel accesible a la conciencia en el cual están mentalmente representadas sólo las propiedades directamente observables. <sup>3</sup>

Las especificaciones del contenido perceptual puro no son intencionales en un grado muy elevado. Al describir la experiencia pura de alguien que contempla la figura de Searle, uno podría decir que "S la vio como dos círculos de 3 milímetros al lado uno del otro y a 4 milímetros el uno del otro con una barra horizontal en lo alto". Esta atribución da resultado cuando se describe la figura de manera tal que el oyente pueda comprender a qué se asemeja la figura; el oyente debe suponer que la experiencia perceptual de S era del tipo de la que se tiene cuando se mira una cosa de ese género. De manera que la descripción fija un estado de cosas que entonces puede ser visualmente imaginado. La especificación será buena en la medida en que permita al oyente hacer lo que se ha dicho. Por consiguiente, hay una especie de objetividad en la especificación de las apariencias que las hace semánticamente más próximas a las especificaciones de información.

Es arriesgado inferir (partiendo de hechos relativos a las atribuciones de contenidos) conclusiones referentes a los contenidos mismos. Los filósofos son proclives a tratar de sacar conclusiones tales como "S ve X como F" o "S tiene una experiencia de algo como F". Sin duda la fuerza de estas expresiones está en indicar no sólo el objeto de la percepción sino también el modo de presentación del sujeto; y esta especificación es psicológicamente informativa. Pero sería erróneo considerar la verdad de "Fido vio el objeto como un hueso", lo cual implica que Fido utilizó un *concepto* en su acto perceptual. Hay otras maneras de dar cuenta de la verdad y de la utilidad de semejante enunciación que atribuye a Fido, no un concepto, sino una percepción o un prototipo o una imagen de hueso. La enunciación imputa al perro una representación puramente perceptual que tiene como su "extensión" una clase de objetos semejantes a huesos típicos cuando se los ve desde diversas orientaciones y distancias en un registro de las condiciones de visión caninas. Por supuesto quien hace la atribución utilizando la palabra "hueso" evoca en su oyente humano el concepto de hueso. Pero el oyente comprende que lo que se imputa es una percepción antes que un concepto.<sup>4</sup>

R. Gregory (1980) da una lista de casos en los que las percepciones son como *hipótesis*. Pone de relieve que la percepción no es simplemente una recepción pasiva de la información procedente de los sentidos. La percepción supone una selección activa y una extracción, al mismo tiempo que la selección de una

representación preexistente en el repertorio almacenado de S, representación acorde con la información sensorial de una manera más o menos semejante a aquella en que la hipótesis está fundada en los datos empíricos. Estoy de acuerdo con este punto de vista, que no hace de la percepción un proceso intelectual; Gregory no dice que las percepciones sean hipótesis conceptualizadas sino que dice tan sólo que son como hipótesis.

El diagrama revela de manera bastante clara que el concepto de "dato" (es decir lo que se le da al sujeto en la percepción) es ambiguo. Lo que se le da inicialmente a la persona, considerada como un sistema de procesamiento de la información, es idéntico a lo que es dado por los órganos de los sentidos, a saber, energías físicas que son vehículos de la información. En el acto de percibir, S no tiene conciencia de estas cosas. Los procesos de información que se siguen no le son dados al sistema sino que son producidos por la actividad del sistema. Por otra parte, lo que le es dado fenoménicamente al sujeto, considerado como un espíritu consciente, es el primer producto accesible a la conciencia de ese tratamiento de la información. En este sentido lo dado es una representación perceptual primitiva provista de un contenido mental. Si el dato está estructurado, lo está como resultado de la actividad del sistema representacional perceptual, pero no está estructurado por S o por la voluntad de S. Desde el punto de vista epistemológico, la significación principal del modelo es que existe algo así como lo dado fenomenológico y que éste no está conceptualizado. La interpretación de lo dado es algo que S no puede hacer; pero si lo hace, esto ocurre posteriormente en el proceso de tratamiento o procesamiento de la información.

La tesis de que el pensamiento puro y la percepción pura se producen por obra de distintos sistemas de representación no se ve viciada por la existencia de estados de espíritu "mezclados" o por la existencia de conexiones intersistémicas. Ciertos estados son perceptoconceptuales: S identifica un objeto en virtud de un modo de presentación demostrativo perceptual y juzga que ese objeto tiene cierta propiedad haciendo entrar en juego un concepto (por ejemplo, S mira a un hombre y ve que es francés). Semejante estado no es ni una percepción pura ni un pensamiento puro, pero uno debe poder distinguir el elemento perceptual del elemento conceptual utilizando criterios para referir las dos cosas a sus respectivos sistemas. Los dos elemen-

tos se combinan en esa utilización por medio de una regla que no forma parte de aquellas reglas que definen un sistema particular.

Ciertos conceptos tienen vínculos internos intersistémicos. Por ejemplo, según el análisis propuesto por Mc Ginn (1983) de los conceptos observacionales, poseer el concepto *rojo* exige que S haya tenido una experiencia visual del rojo. En consecuencia un ciego no puede tener el concepto integral del rojo aun cuando pueda mostrar de manera impresionante la aptitud de desplegar colores en ciertos contextos. Para Mc Ginn, los conceptos observacionales están por fuerza fundados perceptualmente. Si esto es así, el sistema conceptual está íntimamente relacionado con el sistema perceptual (y en este sentido el empirismo clásico queda parcialmente rehabilitado), pero esto no quiere decir que los dos sistemas se fundan en uno solo.

E. Rosch (1975) ha sostenido que numerosos conceptos de especies tales como ave, árbol, perro, están íntimamente ligados con prototipos. Se trata de representaciones mentales de instancias típicas de las respectivas especies. Es posible que esos prototipos correspondan a los sistemas perceptuales de representación. Que yo sepa, Rosch no afirma que S debe poseer (por ejemplo) un prototipo de ave para tener un concepto de ave. Sus descubrimientos muestran que la mayor parte de las personas poseen en realidad el prototipo y el concepto si poseen uno de ellos en el caso de un gran número de categorías familiares básicas.

Necesitamos saber más sobre la naturaleza de estos sistemas de representación. Nos gustaría saber, por ejemplo, qué tiene de específico el sistema visual que lo haga distintivamente visual. ¿Funciona ese sistema según principios representacionales diferentes de los principios de los otros sistemas? Aquí no puedo dar una respuesta completa. Un tratamiento adecuado de semejantes cuestiones iría más allá de la psicología ordinaria y debería apelar a la bibliografía científica. Tal vez haya dos dominios distintos de investigación. La psicología experimental y la fisiología estudian las operaciones físicas cumplidas en el SNC (sistema nervioso central), en tanto que la fenomenología analiza las propiedades de las apariencias visuales. Los físicos sostienen que las dos investigaciones deben juntarse porque las representaciones que tienen propiedades de sensación son estados físicos del cerebro; pero los

dualistas rechazarán esta tesis. Sin embargo, las dos partes pueden ponerse de acuerdo (si no adoptan una postura eliminatória en cuanto a las apariencias) en que la ciencia no rechaza los hechos fenomenológicos que, antes bien, son explananda para la ciencia.

El libro reciente de David Marr, *Vision* (1982), ofrece un modelo de tratamiento detallado de la información que, aunque especulativo, postula tres niveles de representaciones visuales: el esbozo primario, el esbozo “dos y medio” (representación de estímulos casi tridimensionales centrados partiendo del sujeto) y el “modelo tridimensional” (3D), que representa toda la información en profundidad y en un sistema de coordenadas centrado en el objeto. El tercer sistema funciona con elementos perceptuales de construcción de formas 3D cerradas, como cilindros, bloques, secciones, conos, líneas. La teoría de Marr es una teoría del sistema periférico subpersonal pero puede armonizar con la fenomenología. Creo que Marr sostiene que el esbozo “dos y medio” y el modelo 3D son accesibles a la conciencia de S.

El saber común sostiene que los pensamientos son como frases mientras que las percepciones visuales son como imágenes. ¿Cuáles son exactamente las diferencias que hay entre frases e imágenes? ¿Y qué decir de los otros medios o estilos de representación tales como iconos, mapas, diagramas, esquemas, modelos, dibujos, esculturas, etc.? Numerosos filósofos han analizado las características de los sistemas de símbolos. El estudio independiente de las diferencias que hay entre los sistemas pictóricos y los sistemas lingüísticos podría arrojar luz sobre las diferencias entre los sistemas visuales y los sistemas conceptuales y podría suministrar ciertos términos descriptivos útiles. Por ejemplo puede ser esclarecedor preguntarse si la distinción “análogo-digital” puede aplicarse al dominio mental.

Todo pensamiento y toda experiencia visual están estructurados. Por ejemplo, el pensamiento singular de que Reagan está bronceado tiene algo de análogo en la articulación sujeto-predicado pues está compuesto de un concepto *Reagan* como sujeto y de un concepto *bronceado* como predicado, conceptos que están combinados de una manera específica (la predicación) que permite al resultado tener una condición de verdad. Todo S que posee el concepto *Reagan* tiene *ipso facto* la

capacidad de emplear en combinación ese concepto con cualquiera de los otros conceptos de su repertorio para formar otro juicio sobre Reagan. Asimismo poseer el concepto *bronceado* implica la aptitud de combinarlo con cualquiera de los conceptos individualizantes que puedan ocupar una posición de sujeto. Gareth Evans ha sostenido que el pensamiento implica necesariamente el ejercicio compuesto de por lo menos dos aptitudes conceptuales separadas y llamó a esta operación “la condición de generalidad” (Evans 1982, págs. 100-105 y sobre todo la nota 22 de la pág. 104).<sup>5</sup>

La impresión visual más simple, por ejemplo, la de una mancha negra en una superficie blanca, está fenoménicamente estructurada porque articula lo que se ve en la figura y en el fondo. También hay aquí algo así como una “condición de generalidad”. Si S puede tener esta experiencia visual, debe tener también la capacidad de ver una mancha blanca en diferentes posiciones sobre un fondo negro; y si S puede distinguir colores, será capaz de ver diversas manchas de color sobre un fondo blanco (si los colores están dentro del foco de visión). Se puede sostener la idea de que esta aptitud de ver diferentes figuras sobre diferentes fondos es esencial en la experiencia visual humana. Hasta la impresión visual de una superficie uniformemente blanca tiene partes (en virtud de una articulación forma/fondo) y está estructurada. Y esto es así porque la apariencia visual se produce dentro del espacio visual y porque todo espacio visual puede subdividirse en regiones. Verdad es que no se presenta ninguna línea divisoria real entre regiones espaciales, pero el lado izquierdo se verá como diferente del lado derecho aunque no haya línea de separación. Si S tuviera que prestar atención al lado izquierdo o al lado derecho, S manifestaría una estructura de partes espaciales que se encuentra ya presente en la experiencia. Y por supuesto, la mayor parte de la experiencia visual está mucho más estructurada que estos casos mencionados.<sup>6</sup>

Si los diversos sistemas de representación están compuestos de elementos dispuestos en estructuras, ¿cuáles son esos elementos? ¿Tiene cada sistema su propio caudal de elementos, su propio “vocabulario”? ¿Y tiene sus propias maneras de combinarlos con miras a formar complejos?

En el nivel de los pensamientos y también en el nivel de las percepciones es posible trabajar con la hipótesis de una distin-

ción entre simple y complejo, pero hay diferencias entre los dos niveles.

Un sujeto que posee el concepto de *rojo* y el concepto de *casa* puede sintetizar y formar el concepto complejo *casa roja*. Para él el concepto es complejo por el hecho de que se presenta él mismo como complejo, pues es introspectivamente analizable en dos componentes. Para otro sujeto que adquiere un concepto C después de haber estado sometido a la experiencia de casas rojas, C puede denotar casas rojas sin dejar de ser subjetivamente simple. Ese sujeto no piensa en casas rojas de manera compuesta, sino que piensa en ellas como unidades indisolubles de formas y de color. En un caso, el concepto tiene una complejidad interna que S puede encontrar inmediatamente utilizando el concepto. En el otro caso, el concepto no tiene ese carácter compuesto. Sin embargo, en ambos casos S siempre podrá (mediante la reflexión) encontrar algunas de las conexiones lógicas que posee el concepto. De manera que el concepto simple y el concepto complejo de casa roja implican asimismo el concepto de *construcción*. Una especie de análisis revela a S en la introspección esta propiedad lógica. Así, en este sentido de la palabra “análisis”, todo concepto es analizable, hasta aquellos que son subjetivamente simples, porque todo concepto está lógicamente relacionado con otros. La identidad de un concepto dado está fijada no sólo por su complejidad interna (por supuesto, no es esto lo que fija la identidad de los conceptos simples), sino también por su posición dentro de una red. Porque el “vocabulario” conceptual de S es holístico en ese sentido, el repertorio de pensamientos posibles de S es igualmente holístico. Para tener un pensamiento, S debe poseer la capacidad de tener muchos otros. Si S no comprende que un padre es un hombre, luego S no tiene el concepto (completo) de *padre*. Sin embargo la red no flota enteramente en el aire, porque los conceptos observacionales están fundados en las percepciones y serán empíricamente aplicados.

Consideraciones paralelas pueden aplicarse al sistema visual. Si S tiene almacenado un elemento de la forma de un bloque de piedra y un elemento de la forma de una pirámide, puede imaginar la apariencia de un techo colocado encima de cuatro paredes. Puede dejar un espacio entre el techo y las paredes, como si éstas hubieran sido llevadas por el viento, o puede colocarlas de manera adyacente una en relación con la

otra, pero verlas todavía como dos partes distintas. Por otro lado, el mismo S puede tener una representación visual de un objeto con la forma de una casa sin percibir el techo como algo separado de las paredes y ver el conjunto como una simple figura compuesta de partes. No obstante, toda representación perceptual que parece simple a primera vista puede ulteriormente hacerse compleja para S o puede ser reemplazada por una representación compleja que tenga la misma forma fenoménica de conjunto. Para terminar, quisiera considerar la cuestión de saber si la tradicional distinción entre representaciones “análogas” y “digitales” puede aplicarse al dominio de lo mental. Las similitudes y las diferencias apuntadas hasta aquí no resuelven esta cuestión, no son pertinentes para responder a ella. Me referiré a la evaluación que hace Goodman de la distinción análogo / digital en *Languages of art* (1968) cap. IV, que, en mi opinión, representa lo esencial de la concepción tradicional.

Goodman distingue un *esquema* simbólico (un conjunto de caracteres que poseen propiedades sintácticas y relaciones) y un *sistema* de símbolos (un esquema cuyos caracteres tienen relaciones semánticas —denotativas— con clases de objetos). Goodman define “análogo” y “digital” primero en el caso de los esquemas y luego en el de los sistemas.

Un esquema es análogo si es sintácticamente denso. Considérese un dispositivo con una esfera redonda [como de reloj] y un señalador que se detiene en toda orientación de la esfera. Todo punto de la circunferencia es un “carácter”. Ese continuo de caracteres es sintácticamente denso: “ofrece una infinidad de caracteres ordenados de manera tal que entre dos de ellos hay siempre un tercero” (pág. 316). Un esquema en cambio es digital si sus caracteres están diferenciados.

Es posible determinar en todo ejemplar de carácter (*token*) a qué tipo de carácter pertenece (en el interior de ese esquema).

Un sistema es análogo si utiliza un esquema análogo que sea semánticamente denso (págs. 152-153). Ejemplos serían aquellos de sistemas cuyos caracteres denotan continuamente magnitudes de cantidades o grados de cantidades continuamente variables. La esfera y el señalador, dispuestos de manera que varíen con la presión del aire, formarían pues un sistema completamente análogo. Para poder ser un sistema digital, un esquema digital debe estar semánticamente diferenciado, lo

cual quiere decir que, habiendo dos caracteres, si un objeto es denotado por uno u otro de los caracteres pero no por los dos, debe ser posible determinar cuál es el carácter que denota ese objeto (véanse págs. 152-153). La extensión de los símbolos no debe ser vaga.

Las representaciones mentales tienen a la vez una dimensión sintáctica y una dimensión semántica, como todas las representaciones. De manera que debemos responder a dos pares de preguntas. Consideremos en primer lugar los conceptos. ¿Pertenecen los conceptos a un esquema digital? Esto depende de lo que sean los caracteres del esquema. Supongamos que cada concepto sea un carácter y que cada uno de ellos se diferencie de todo otro concepto por su función inferencial; supongamos también que en cada utilización de un concepto particular (*token*), el concepto tipo a que pertenece esté determinado. Entonces el esquema conceptual está sintácticamente diferenciado y por lo tanto es digital.

La cuestión de saber si constituye un *sistema digital* es extremadamente dudosa, pues en su conjunto no parece estar semánticamente diferenciado. Podrían existir ciertos pares de conceptos, y normalmente deberían existir, de tal condición que existan objetos denotados por uno u otro de los caracteres y no por los dos, pero es imposible determinar cuáles. Hay objetos que no entran claramente dentro del concepto *pesa una onza o más* ni claramente en el concepto *pesa menos de una onza*, por más que sepamos cuáles objetos deben entrar en uno o en el otro concepto, puesto que queda agotado el dominio de aplicación del concepto. Otro ejemplo podría ser el par de conceptos *rojo* y *anaranjado*. Por supuesto, hay conjuntos de conceptos que dividen un dominio en clases las cuales tienen límites absolutamente claros, como la familia de conceptos *pesa exactamente el peso de una moneda de un franco, contiene dos francos, contiene tres francos*, etc. Semejantes familias crean subsistemas digitales dentro de un sistema que no es digital en su conjunto. El sistema conceptual tiene un papel principal: nos permite representar todos los aspectos de la realidad en todas las maneras que nos son útiles. Por ejemplo, la capacidad que tiene el sistema conceptual de representar a la vez cantidades discretas y cantidades continuas, a la vez vaga y precisamente, y esto es lo que cabe esperar.

Si consideramos ahora las percepciones, debemos preguntarnos primero si los caracteres forman un esquema sintácticamente denso. Considérese la percepción visual de una línea recta sobre un fondo blanco. La apariencia visual de la línea puede subdividirse en el interior del espacio visual. Supongamos que cada punto de la línea representa un carácter. En ese caso el conjunto de los caracteres que forman la línea constituye un conjunto denso. El esquema *es* análogo.

Un objeto así representado, como por ejemplo un trozo de cable eléctrico, puede subdividirse en partes más pequeñas aunque no haya líneas divisorias preexistentes. De suerte que la percepción está compuesta por un sistema de caracteres semánticamente denso. El mundo de las apariencias visuales imita los continuos de la realidad al crear un espacio mental en el cual existen continuos aparentes. Todo sistema que explota de esta manera las propiedades espaciales de las representaciones es necesariamente análogo.

Compárese esto con un “dispositivo de reconocimiento de formas” (*pattern recognizer*) que contiene un vocabulario de patrones en los que pueden repartirse las informaciones. Supongamos que ese vocabulario esté constituido por cuatro formas (círculo, cuadrado, triángulo, hexágono), cada una de ellas de diez tamaños posibles. Su esquema de símbolos es digital, pero ¿está semánticamente diferenciado? Desde un punto de vista objetivo hay casos límites, hay formas que no son ni claramente esferas ni claramente no esferas. Pero si especificáramos un nivel de precisión de la resolución, entonces los símbolos estarían semánticamente diferenciados.

Hasta en este caso puede haber un objeto de tal condición que los mecanismos no puedan decidir (en un ensayo) qué patrones seleccionar (por ejemplo, la esfera de tamaño cuatro o la esfera de tamaño cinco) porque el objeto está demasiado lejos o es confuso. El mecanismo no puede dar una respuesta determinada; el hecho de que el objeto entre en una representación o en otra es un hecho perfectamente determinado.

Esto muestra que la palabra “determinado” es ambigua. Tiene un sentido matemático y un sentido práctico. Goodman trabaja con la noción formal de “determinar si un objeto entra en la clase de conformidad con un símbolo”; pero aquí estamos hablando de un mecanismo *que aplica* un símbolo a una clase.

Aquí tenemos necesidad del concepto de adaptación o de aproximación. Un huevo no es perfectamente oval para que un sujeto S le aplique su representación oval. ¿Sabe S que el huevo no es perfectamente oval? S se lo representa como un óvalo, pero la retina de S recibe la información de que el huevo no es oval. Centros de procesamiento de la información visual de los más elevados en la jerarquía han simplificado o ignorado la información porque lo que les interesa es la mejor aproximación.

Nuestra conclusión sostiene que el sistema conceptual no es íntegramente digital, pero que el sistema perceptual es análogo en ciertas dimensiones cualitativas. Para llegar a este resultado hemos imaginado varias hipótesis artificiales y la más dudosa de ellas es la de suponer que sabemos bastante de las representaciones mentales mismas para decir cuáles son sus caracteres y si están sintácticamente diferenciadas o son densas. En el caso de los conceptos, hemos supuesto en efecto que existe un lenguaje del pensamiento. Porque supusimos que cada concepto tiene su propio perfil de propiedades funcionales, hemos supuesto también que cada concepto debe diferir de los demás por su naturaleza intrínseca. Sin embargo, en el caso de las representaciones perceptuales hemos sustentado una hipótesis suplementaria: la de que las propiedades fenoménicas de dichas representaciones nos guían para comprender su naturaleza intrínseca (“la apariencia de una línea es continua y en consecuencia la representación subyacente es continua”). Pero, como lo sabemos por consideraciones teóricas y por la familiaridad que tenemos con los dibujos que hacen los ordenadores, es posible simular una imagen en un dispositivo digital en un grado deseable de resolución de la imagen. Lo que aparece en la pantalla es realmente digital, pero ofrece la misma apariencia que tendría si fuera análogo. *Pari passu*, pudiera ser que el mecanismo subyacente de la visión sea en realidad digital, por más que produzca experiencias que *parecen* solamente análogas. Se ha dicho con frecuencia que el cerebro es un ordenador digital. Quienes lo dicen tal vez sean víctimas de una confusión sobre el sentido de “digital” y tal vez tengan una visión simplista de la transmisión neurónica. Sea ello lo que se quiera, el problema permanece en pie y yo me contento con dejarlo en ese estado, consciente de no haber más que arañado su superficie. Se presentan dificultades cuando uno extrapola partiendo de datos fenomenológicos para llegar a conclusiones relativas a la

naturaleza de los sistemas de representación; pero esas dificultades no amenazan la tesis (fundada en diversos datos) de que realmente hay sistemas distintos de representación.

## Notas

<sup>1</sup> Es posible que hoy Churchland ya no sostenga este punto de vista. En un artículo más reciente (Churchland 1985), afirma que un sujeto ciego o incapaz de percibir los colores no tiene lo que tiene un sujeto visualmente normal, es decir, “una representación del rojo en un medio prelingüístico o sublingüístico de representación en el caso de variables sensoriales, ni la capacidad de hacer discriminaciones sensoriales”. Sin embargo Churchland puede hacer referencia aquí a un medio para representar cualidades o propiedades en el mundo exterior. La distinción tradicional entre la sensación y la percepción es una división entre estados que son y estados que no son representacionales, o entre estados considerados como no representacionales y los mismos estados considerados como representacionales, según la expresión de Hamlyn (1961). Yo prefiero la concepción de Peacocke de que las propiedades sensoriales están presentes en los mismos estados que aquellos que poseen propiedades representacionales. Sin embargo, a diferencia de Peacocke, sostengo que esas representaciones del mundo (en sus formas puras y no primitivas) son no conceptuales.

<sup>2</sup> La etapa ilegítima del razonamiento se encuentra en la pág. 44 de Strawson 1979: “No podemos dar una caracterización verídica ni siquiera de la experiencia sensible más allá de la cual van esos juicios (según la expresión de Ayer) sin hacer referencia a esos juicios mismos. Y hasta nuestra experiencia consciente está fuertemente impregnada de conceptos de los objetos que figuran en esos juicios”.

<sup>3</sup> Sobre este particular y paralelamente a lo que dice Peacocke, pág. 17, en relación con la imagen del pato-conejo, se puede observar que el objeto de la percepción es aquí él mismo una representación, un símbolo pictórico o lingüístico. El hecho de que aparezca en la página de un libro no puede dejar de ser advertido por el observador. Las experiencias que describe Searle están todas construidas atendiendo a la capacidad que tiene el observador de ver, a través de los rasgos trazados con tinta, los posibles objetos que ellos representan; y el observador entrenado tendrá disposición a decir que percibe esos objetos en la figura. Las percepciones de estados de cosas ordinarias ofrecen mejores ejemplos. Pero aun en esto, adultos experimentados no se fían de las apariencias y a menudo van directamente a las interpretaciones.

<sup>4</sup> Peacocke (1983) escribe: “Es propio de la naturaleza del contenido representacional que éste no pueda construirse partiendo de conceptos a menos que el propio sujeto de la experiencia tenga los conceptos en cuestión: el contenido representacional es la manera en que la experiencia presenta el mundo y difícilmente puede presentar el mundo de esa manera si el sujeto es

incapaz de apreciar lo que ella es. Únicamente quienes disponen del concepto de esfera pueden tener la experiencia de una esfera que esté ante ellos y únicamente quienes disponen de conceptos espaciales pueden tener experiencias que representen las cosas como distribuidas en profundidad dentro del espacio.” (pág. 7) Lo que aquí afirma Peacocke es compatible con la posición del autor de este artículo, pues aquél no sostiene aquí la tesis general de que en el caso de todo F, la experiencia que tiene S de algo como F requiere que S posea el concepto F. Más adelante (pág. 19), Peacocke parece adoptar la tesis general. Por mi parte critico la tesis general. Admito que ciertas percepciones requieren el dominio de conceptos específicos, pero esas percepciones no son primitivas. Peacocke da conceptos como el de *esfera* y conceptos de relaciones espaciales en profundidad. Estos pueden ser especiales porque Peacocke los considera conceptos primitivos *teóricos*. Enuncia la tesis de que el niño durante su desarrollo construye una teoría del mundo objetivo y que las experiencias espaciales no son accesibles a un niño que no posea la teoría. Sobre este punto particular creo que Peacocke se equivoca. Los datos que poseemos sobre el desarrollo sugieren que los niños se representan figuras tridimensionales y relaciones tales como “delante”, “detrás”, “más alejado que” de manera puramente perceptual, antes de adquirir los conceptos correspondientes (Bower 1974). Los niños aprecian esos rasgos del mundo ejerciendo facultades perceptuales que probablemente son innatas. Por supuesto, esto no significa negar que tengan necesidad de cierto tiempo y de cierta cantidad de experiencia desencadenante antes de poseer su dispositivo completo.

<sup>5</sup> Las tesis de Evans no implican que éste sostenga que los conceptos son símbolos dentro de un lenguaje del pensamiento. Por eso Evans prefiere hablar de aptitudes.

<sup>6</sup> Sobre prudentes observaciones relativas a la división de las percepciones en partes, véase Wittgenstein (1953), sección 47. Wittgenstein parece negar que haya una manera científica de hacer esta división porque la distinción parte/todo es relativa a un interés particular (es decir, puede figurar en diferentes juegos de lenguaje). Este autor dice: “A la cuestión filosófica: ‘¿Es compuesta la imagen visual de este árbol y cuáles son sus partes componentes?’, la respuesta correcta es: ‘Todo depende de lo que se entienda por compuesto’”. (Por cierto, ésta no es una respuesta sino que es un rechazo de la cuestión.) Aquí “filosófico” quiere decir, según presumo, “fuera del contexto de un objetivo práctico”. Por mi parte, por supuesto, considero la cuestión dentro del contexto de la busca de una respuesta verdadera y científica.



## Referencias bibliográficas

- D.M. Armstrong 1963 *Perception and the physical world*, Routledge, London.
- T.G.R. Bower 1974 *Development in infancy*, Freeman, San Francisco.
- P.M. Churchland 1979 *Scientific realism and the plasticity of mind*, Cambridge University Press.
- P.P. Churchland 1985 "Reduction, qualia, and the direct introspection of Brain states", *Journal of Philosophy*, LXXXII, enero
- G. Evans 1982 *The varieties of reference*, Clarendon, Oxford.
- J. Fodor 1983 *The modularity of mind*, MIT Press (tr. fr. París Minuit 1985).
- J. Fodor 1985 "Précis of *The modularity of mind* with replies to peers commentaries", *Behavioural and Brain sciences*, 8,1.
- K. Foster 1985 "Commentary on Fodor", *Behavioural and Brain Sciences*, 8,1.
- N. Goodman 1968 *Languages of art*, Bobbs Merrill, Indianapolis.
- R. Gregory 1980 "Perceptions as hypotheses", *Phil. Transactions of the Royal Society of London*, B, 290, 181-197.
- D.W. Hamlyn 1961 *Sensation and perception*, Routledge, Londres.
- J. Hintikka 1975 "The intentions of intentionality", cap. 10 de *The intentions of intentionality*, Dordrecht, Reidel.
- D. Marr 1982 *Vision*, Freeman, San Francisco.
- M. Ginn 1983 *The subjective view*, Clarendon, Oxford
- V. Neisser 1983 *Cognitive psychology*, Nueva York
- C. Peacocke 1983 *Sense and content*, Clarendon, Oxford
- E. Rosch 1975 "Cognitive representations of semantic categories", *J. of exp. Psychology, General*, 104, 192-233.
- G. Sperling 1960 "The information available in brief visual presentations" *Psychological monographs*, 74, 11.
- P.F. Strawson 1978 "Perception and its objects", en *Perception and identity*, G. McDonald, (comp.), Blackwell Oxford.
- L. Wittgenstein 1953 *Philosophical Investigations*, Blackwell, Oxford.

## 3

# Psicología del sentido común y ciencias cognitivas

Samuel Guttenplan\*

## 1. ¿Qué hay en un nombre? No pocas cosas

No me gusta la designación "psicología ordinaria". Me parece que esta expresión implica falsamente que la psicología ordinaria es una empresa mucho más unitaria y sistemática de lo que es en realidad; pero no quiero tratar aquí este punto. Me propongo más bien criticar la palabra "ordinaria" que figura en el rótulo. En particular creo que esa palabra nos induce a seguir dos direcciones que son problemáticas. Una manera de comprender esa palabra implica que la psicología ordinaria es simplemente una parte del conocimiento no reflexivo común de que están provistos los seres humanos, es decir, un esquema que todos utilizamos en nuestros ingenuos intentos de explicar la conducta de nuestros semejantes humanos. La otra manera de comprender ese término es suponer que la psicología ordinaria es algo semejante a la medicina común. Es el producto, lleno de defectos pero reflexivo, de generaciones de hombres cuya comprensión del alma humana no es ya más exacta que la de generaciones de profanos que prodigaron sus opiniones sobre la salud del cuerpo sin un conocimiento real de la ciencia biológica.

Me parece que no debemos echar a andar por ninguna de estas dos sendas por razones complejas que están estrechamente ligadas. Y al resistir a estas tentaciones podremos aprender algo sobre las relaciones que pueda haber entre la psicología ordinaria y la ciencia cognitiva. De manera que la resistencia a la tentación es el tema de este artículo.

\* © Samuel Guttenplan, 1988.

En la edición original francesa, este artículo fue traducido del inglés por Pascal Engel [E.]

## 2. Psicología ordinaria y física ingenua

La primera manera de comprender la designación “psicología ordinaria” alienta la idea de que ella es como la física del sentido común: el conjunto de previsiones comunes que rige de manera irreflexiva nuestra concepción del mundo de objetos de tamaño mediano. Un artículo reciente de Andy Clark nos suministra una buena fuente para esclarecer este género de concepción (1987). Veamos una enunciación más completa de la tesis:

“¿No es la psicología del sentido común la defectuosa especulación sobre lo mental del hombre de la calle? No lo creo. Para comprender por qué, puede ser útil trazar un paralelo con la concepción de la física ingenua de Hayes (1987). Una física ingenua es un cuerpo de conocimientos comunes sobre leyes físicas y conceptos que nos ayudan a desenvolvernos en nuestro mundo de objetos macroscópicos de todos los días. Cierta clase de conocimientos (no necesariamente formulados de manera lingüística) sobre conceptos y relaciones como ‘fluido’, ‘causas’, ‘sostén’, ‘encima’, ‘debajo’; y ‘al lado de’ es vital para un ser móvil que debe manipular objetos.

Lo que es cierto en el caso de una física ingenua debe igualmente serlo (si tengo razón) en el caso de una psicología ingenua. Así como el ser móvil tiene necesidad de saber lo que es un sostén, así también un ser socialmente móvil debe saber algo sobre los estados mentales (creencias, deseos, motivaciones) de sus iguales (págs. 144-5).”

Ahora bien, por más que yo piense que los seres humanos adultos y normales poseen los elementos básicos del esquema —conceptos como los de creencia, deseo, intención, motivación y emoción—, me parece que el esquema difiere del de la física ingenua en dos maneras importantes. En primer lugar, la psicología ordinaria cuando opera realmente de un modo que puede reconocerse como una teoría explicativa está lejos de ser no reflexiva. En segundo lugar, la psicología ordinaria no es —lamentablemente en mi opinión— una posesión común de los seres humanos tanto como lo supone la designación. Precisemos estos puntos:

Que la psicología ordinaria deba o no deba entenderse en última instancia como una *teoría explicativa empírica* (en cualquiera de los sentidos técnicos de este término), lo cierto es que ella opera como si lo fuera. Cuando hay que explicar diversos aspectos del comportamiento humano se apela a ella y nosotros explicamos nuestro mundo social en sus términos. Como ya dije, los intereses centrales típicos del esquema son conceptos como los de creencia, de deseo, de intención, de motivación y de emoción, aunque es importante reconocer que la lista podría ser considerablemente más larga y más variada. Luego volveré a ocuparme de este punto. Y, como también observé antes, esos conceptos evidentemente forman parte del caudal conceptual de todo adulto humano normal. Pero aquello que hace de la psicología ordinaria una clase de *psicología* no es simplemente el hecho de que contenga tales conceptos, sino además el que sean utilizados de la manera en que lo son en la comprensión y descripción del comportamiento de los seres humanos en general y en particular. Y por “utilizados” entiendo que esos conceptos se encuentran entremezclados en análisis explicativos totalmente específicos de la conducta humana, análisis que pueden corregirse o aumentarse bajo la presión de la observación y de otras teorías. Veamos dos ejemplos.

I) Una historia verdadera: Barbara y Richard se conocían desde pequeños y muy pronto se hicieron amigos. Cuando tuvieron alrededor de diecinueve años su amistad se convirtió en algo más profundo, aunque no hubieran sufrido para ello la presión de los vínculos sociales que existían entre sus respectivas familias. Barbara y Richard decidieron que cuando tuvieran veintiún años se casarían, de manera que se pusieron a acariciar ese proyecto con gran placer de los dos grupos de padres. Pero su matrimonio duró sólo un día, es decir que habiéndose casado después de quince años de una amistad profunda pasaron solamente un día juntos en la casa que habían escogido con tanto cuidado para vivir. ¿Por qué?

Aquellos de nosotros que conocían a la pareja habían forjado numerosas teorías sobre la ruptura. Cada una de las teorías que proponíamos utilizaba el registro de conceptos de la psicología ordinaria, aunque las teorías diferían mucho en cuanto a la estructura y a los contenidos de creencias, deseos, etc. que atribuíamos a Barbara y a Richard.

II) La historia de Barbara y de Richard invita a hacer toda clase de especulaciones teóricas sobre lo que había ocurrido. En este caso, las teorías están vinculadas con hechos específicos sobre aquellos dos individuos. Pero la psicología ordinaria interviene también en conceptos más amplios. Cuando en los Estados Unidos comenzaron las recientes audiciones "Iran/Contra", fue universalmente condenada la conducta ilegal de los principales actores de ese asunto. Pero cuando North prestó testimonio, el coro de desaprobaciones se volvió bruscamente contra aquellos que lo interrogaban. En los Estados Unidos y fuera de los Estados Unidos, a muchos les pareció aquello muy extraño. Para los comentaristas políticos el enigma estaba en saber cómo el autor de malas acciones confesadas podía llegar a ser tratado como un héroe por los mismos grupos que antes se habían mostrado tan dispuestos a condenar sus fechorías.

Como en el primer caso, era posible elegir dentro de una vasta gama de teorías particulares. Cada una de ellas atribuía actitudes y emociones realmente diferentes a North y a diversos sectores del público y de la prensa. Algunos hasta apelaban a las concepciones de las raíces históricas de la sociedad norteamericana y a sus rastros que estaban presentes en la trama social.

En cada uno de estos dos ejemplos nos hallamos en presencia de un fenómeno que pide una explicación y en presencia de todo un conjunto de teorías al respecto. Lo que las teorías tienen en común es, desde luego, que utilizan materiales tomados de la estructura de la psicología ordinaria. Volveré a tratar este punto y por el momento dejaré deliberadamente en la vaguedad la naturaleza de esos materiales y de esa estructura. Desde este punto de vista, las teorías de la psicología ordinaria son mucho más efímeras que muchas teorías científicas. Con frecuencia se las formula súbitamente y se las refuta con la misma rapidez. Por ejemplo, fue rechazada muy pronto la sugestión de que Barbara había tenido una primera experiencia sexual desastrosa con Richard en su noche de bodas. Barbara y Richard se acostaban con frecuencia desde hacía mucho tiempo y aquel día, como cabía esperar, se encontraban demasiado fatigados para entregarse a aquellas relaciones después del casamiento. Además, ¿cómo una experiencia negativa entre personas que habían sido amigas durante quince años podía bastar para impedir por lo menos un breve período de prueba para ver si el matrimonio marcharía o no? La partida de Barbara al cabo de un día es el

núcleo del misterio, y la teoría sexual no podía explicarlo.

Ante ejemplos de la vida real como este, resulta difícil sostener la idea de que la psicología ordinaria es un saber común no reflexivo. El contraste entre éste y la física ingenua es acusado. Se requiere un gran esfuerzo para comprender algunas de las cosas que experimentan nuestros semejantes. Además, no todos somos igualmente perspicaces; están aquellos que no tienen ningún talento para utilizar teorías psicológicas ordinarias que expliquen la conducta; están aquellos capaces de ver lo que falla en cualquier teoría y están aquellos otros que no entienden la teoría si se la aplica a ellos mismos. Si las diferencias individuales que hay entre nosotros como psicólogos ordinarios se dieran también en lo tocante a nuestras facultades de físicos ingenuos, entonces nuestros contactos con el mundo de los objetos habituales estarían llenos de peligros.

El hecho de que todo el mundo considere las teorías de la psicología ordinaria como irreflexivas se debe en parte a los ejemplos empobrecedores que los filósofos habitualmente aducen en sus exposiciones. Ciertamente, cuando el cartero evita entrar en un jardín en el que hay un gran perro lobo que ladra y gruñe, habrá consenso en cuanto a explicar la conducta del hombre. No se necesita pensar mucho para llegar a tal consenso. Pero casos como este no agotan, nada de eso, este terreno. Recordando hasta qué punto puede ser difícil explicar las cuestiones del mundo humano, los anteriores ejemplos deberían permitirnos apreciar la variedad de talentos que son frecuentemente necesarios para hacer acertadas algunas explicaciones de la psicología ordinaria. Por cierto, todos podemos explicar por qué el cartero no entra en el jardín, pero ese es más o menos el nivel a que puede llegar la mayor parte de los seres humanos. Verdaderos progresos en las teorías sobre los seres humanos exigen una considerable competencia. Algunos escritores e historiadores y aquellos de nosotros que tienen la suficiente imaginación para encontrar explicaciones de la conducta humana cuando las explicaciones están lejos de ser evidentes, son en verdad expertos en psicología ordinaria. En cambio no me parece tener sentido hablar de "expertos" en física ingenua. Cuando alguien comienza a reflexionar seriamente sobre los conceptos físicos ingenuos y adquiere un conocimiento que falta a los demás, aquél ya no es más un físico ingenuo. Es sencillamente un físico.

Se podría pensar que no he hecho justicia a la concepción

de la física ingenua. Puse énfasis en el hecho de que el empleo de nociones de la psicología ordinaria en teorías psicológicas específicas requiere una reflexión y una habilidad considerables. Pero Clark puede haber querido decir que solamente el marco de conceptos de la psicología ordinaria puede compararse con la física ingenua. En otros términos, su tesis puede ser la de que nuestra posesión de conceptos de la psicología ordinaria es el equivalente evolucionista de nuestra posesión de los conceptos ingenuos de la física, pero que nuestro empleo de atribuciones de creencias y deseos puede requerir algo más que el hecho de ser sencillamente un humano común. Sobre este punto es posible hacer dos observaciones.

En primer lugar ( y este es el punto menos importante), esta manera de ver hace que la analogía entre psicología ordinaria y física ingenua sea menos atractiva de lo que podía parecer a primera vista. Pues nuestra física común no es simplemente el dominio implícito de conceptos pertenecientes a cierta estructura, sino que es además un dominio implícito de principios y de explicaciones que rigen esos conceptos. Si la analogía es realmente sólida habría un paralelo entre el carácter natural de nuestra aptitud de eludir un objeto que se encuentra en nuestro camino y nuestra aptitud de especular sobre la conducta de Barbara. Pero la analogía no es realmente sólida. Dentro de ciertos límites, podemos hacer lo primero muy bien sin reflexionar en ello, pero lo segundo sólo puede ser tratado por alguno de nosotros, pues requiere una reflexión muy sostenida de quien intenta especular sobre esa conducta.

En segundo lugar, aquí me cuesta mucho trabajo separar los conceptos del marco de la psicología ordinaria, por un lado, y las teorías que podemos elaborar a su respecto, por otro. Yo aceptaría seguramente que en numerosos casos los conceptos pueden separarse de sus teorías. Pero no me adhiero a la recomendación de esos cultivadores severos de la teoría que sostienen que retirar un concepto del suelo de su teoría lo mata en todos los casos. Sin embargo son cuestiones realmente importantes las de saber cómo puede llevarse a cabo esa separación y saber qué cantidad de suelo teórico puede quitarse. Ciertamente hay algo que podemos llamar el marco de la psicología ordinaria. Por lo menos ese marco consiste en una lista de los conceptos utilizados en las explicaciones de la psicología ordinaria y particularmente esos conceptos que se llaman actitudes proposicionales, por más que tengo mis dudas

sobre la posibilidad de establecer semejante lista. Pero ¿realmente tiene sentido considerar esos conceptos como conceptos que están en posesión de criaturas poco capaces de emplearlos? Con toda seguridad, la aptitud para utilizarlos en toda una gama de circunstancias es crucial.

Algunos trabajos recientes sugieren que los niños adquieren o desarrollan el concepto de creencia entre los tres y los cinco años de edad (Wimmer y Perner 1983). Quizá Clark pensaba en ese conocimiento que, según la opinión de todo el mundo, es irreflexivo. Sin embargo, sin poner en duda la información comunicada sobre estas experiencias, quisiera yo sugerir que lo que adquieren los niños de cinco años es solamente un concepto parcial; lo reconocemos como una creencia porque sabemos que ésa es la primera etapa del desarrollo de algo más rico.

La manera en que el niño de cinco años posee el concepto de creencia es ciertamente irreflexiva, pero esto no nos permite en modo alguno aceptar la analogía propuesta por Clark. Pues los niños de cinco años tienen su manera de crecer. Y cuando crecen van utilizando la psicología ordinaria de manera más reflexiva. Algunos de ellos hasta pueden llegar a ser expertos en psicología ordinaria. Sin esa maduración, no veo cómo podríamos considerar que el niño posee el concepto de creencia. En comparación, aquellos de nosotros que no se preocupan por aprender la física propiamente dicha continúan siendo físicos ingenuos durante toda su vida; sin pensar en lo que hacen pasan por el umbral de las puertas o eluden diversos obstáculos.

Por más que admita yo que los conceptos de la psicología ordinaria puedan separarse de sus empleos particulares cuando se trata de explicar, por ejemplo, por qué Barbara abandonó a Richard, me siento inclinado a pensar que la analogía con la física ingenua implica una separación demasiado acusada. Un concepto llamado "creencia" pero divorciado de su empleo en la psicología ordinaria puede ser algo que poseemos de manera irreflexiva, pero es demasiado endeble para que lo tomemos seriamente por una creencia. Volveré a ocuparme enseguida de estos problemas. Pero primero quisiera ver cómo los ejemplos y consideraciones anteriores pueden aplicarse en otra interpretación del rótulo "psicología ordinaria", una interpretación según la cual puede haber una analogía entre la psicología ordinaria y la medicina común.

### 3. Psicología ordinaria y medicina común

Churchland ataca la psicología ordinaria en estos términos:

“La psicología ordinaria de los griegos es esencialmente la psicología ordinaria que utilizamos hoy, y nuestra aptitud para explicar la conducta humana en sus términos es apenas un poco mejor de lo que podía ser la aptitud de Sófocles. Es un período muy largo de estancamiento y de infecundidad en una teoría” (Churchland 1981).

Y Stich escribe:

“El hecho mismo de que [la psicología ordinaria] sea una teoría *ordinaria* debería hacernos concebir sospechas. Pues en cualquier dominio en el que se pueda pensar, los antiguos pastores y los conductores de caravanas cuyas especulaciones estaban relacionadas con teorías populares eran notoriamente poco fiables. La astronomía popular era una astronomía falsa y no solamente en los detalles. Por maravilloso e imaginativo que fuera el poder popular de teorización, se ha revelado falso en todos los dominios en los que hoy tenemos una ciencia razonablemente ajustada.” (Stich 1983).

Según una de las maneras de entender lo que pueda ser una teoría psicológica ordinaria, estas observaciones tienen poco sentido; sencillamente dejan a un lado la cuestión. Y ahora quisiera explicar por qué. Según otra interpretación, tienen cierto sentido, aunque me propongo explicar en lo restante de este artículo por qué dichas observaciones son falsas. Por el momento lo que hay que recordar es que una teoría de psicología ordinaria es una teoría elaborada con diversos materiales conceptuales a fin de explicar ciertos aspectos de la conducta humana. Las teorías sobre las razones que movieron a Barbara a abandonar a Richard o sobre las razones de la sorprendente reacción suscitada por el testimonio de North son teorías psicológicas ordinarias en este sentido.

Y en este aspecto, los ataques de Churchland parecen absurdos. Lejos de estar estancadas y ser infecundas las teorías psicológicas populares son por el contrario demasiado fértiles.

Las teorías sobre las razones que impulsaron a Barbara a abandonar a Richard no valían nada y la mayor parte de ellas quedaron invalidadas apenas se las confrontó con los hechos. Es claro que Churchland debe de haber querido decir otra cosa cuando se refirió a los dos millares de años de estancamiento. Asimismo, las observaciones de Stich no pueden referirse a las teorías psicológicas ordinarias. Los pastores y los conductores de caravanas de milenios atrás no habrían podido decir nada verdadero o falso sobre Barbara y Richard.

Aquello a lo que se refieren Churchland y Stich no deben ser tanto las teorías particulares con las que explicamos la conducta humana como el marco general en que dichas teorías son explicadas. El lector hasta puede pensar que había cierta perversidad de mi parte cuando consideré otra interpretación, aunque quiero defenderme sobre dos puntos: en primer lugar, al obligar al lector a considerar que la posición de Churchland y Stich puede aplicarse aun a casos particulares, yo quisiera inducirlo a sospechar que las ideas de estos autores sobre la psicología ordinaria son muy extrañas. Me parece enteramente sensato declarar que en lo tocante a la explicación psicológica habitual de la conducta humana estamos en mejores condiciones que los conductores de caravanas y aun que Sófocles en sus explicaciones psicológicas de la conducta y que toda posición que no tome seriamente esa posibilidad me parece sospechosa. En segundo lugar, creo que la cuestión del marco de la psicología ordinaria es la más importante si verdaderamente queremos comprender la relación que hay entre psicología ordinaria y ciencias cognitivas. Mi interpretación perversa de las citas hechas está destinada a obligar a los defensores de Stich y de Churchland a salir al descubierto.

Podemos presumir que la idea general de estos autores es más o menos la siguiente. Las primeras teorías médicas tenían como marco de fondo todo un análisis de los diversos humores exhalados por el cuerpo. Las teorías particulares sobre enfermedades específicas estaban fundadas dentro de ese marco. Y sin embargo sabemos que el marco de conjunto era falso: no había humores y toda teoría fundada en ellos es falaz. Y lo mismo ocurriría en psicología. Quizá los conceptos que forman el marco de la psicología ordinaria no sean mejores que conceptos como el de humor. Tal vez lo que aquí es crucial es la manera en que se especifica el marco de fondo y su relación con las

teorías particulares que hacen el verdadero trabajo de explicación. No es en modo alguno evidente que la analogía médica sea la única posible y muy bien puede ocurrir que sea este un modelo que no resulta plausible para comprender la psicología ordinaria. Pero no se puede discutir esta cuestión mientras no se tenga una idea más clara de las clases de modelos posibles.

#### 4. Marcos y teorías

He de sostener que existe toda una gama de casos en los que algo que podemos llamar un marco teórico entra en interacción con teorías explicativas específicas que están en el interior de ese marco. Veamos tres ejemplos que luego discutiremos:

1) En la genética contemporánea, la teoría de fondo es la biología molecular. Más específicamente, en este caso el marco está representado por la teoría bioquímica de sustancias tales como el ADN, el ARN y esas enzimas proteínicas que tienen la función y la estructura 3-D de los ácidos nucleicos. Teorías genéticas específicas se apoyan en el fondo bioquímico para explicar diversas relaciones entre padres e hijos y la distribución de los caracteres fenotípicos de las poblaciones. Por ejemplo, el concepto de una secuencia de ácidos aminados en una molécula de ADN desempeña una parte decisiva en las teorías particulares de la herencia genética que fueron propuestas en años recientes. Esa secuencia desempeña el papel de marco o fondo de la teoría.

2) El marco de fondo del estudio de la materia incluye, entre otras cosas, la hipótesis de que la materia está hecha de partes cada vez más pequeñas pero que el descenso por la escala tiene un límite finito. En una palabra, podemos decir que dentro del marco de fondo de la física está el concepto de átomo. Por supuesto, la manera en que ese concepto se utiliza en la física contemporánea es diferente de la manera en que lo utilizaban Demócrito o Dalton.

3) Consideremos ahora el estudio teórico del lenguaje, el estudio de las emisiones lingüísticas de los seres

humanos. Con relación a esa masa de hechos, los conceptos de, por ejemplo, frase, nombre sustantivo y verbo son conceptos teóricos y son centrales en las teorías del lenguaje que se llaman las gramáticas. Por supuesto, diferentes gramáticas tendrán diferentes cosas que decir sobre el funcionamiento específico de tales categorías, pero es fácil reconocer cierta constancia en esos conceptos en el fondo de este proscenio cambiante de la escena. Los conceptos de frase, de nombre y de verbo son aproximadamente como el de átomo desde este punto de vista, pues parece que somos capaces de discernir la idea de un átomo en el plano de fondo mismo y relacionarla con los diferentes cambios que han conducido a la física contemporánea.

Sin duda alguna se podrían agregar muchos otros ejemplos; espero sin embargo que éstos permitan aprehender una estructura dentro de lo que es un continuo. En uno de los extremos de la escala están los marcos teóricos que enlazan sus principales conceptos entre sí en virtud de regularidades y de relaciones. Desde este punto de vista, toda teoría explicativa particular que emplee los conceptos pertenecientes a ese marco se verá obligada a respetar esos enlaces. Si se utiliza un concepto de manera que se violen esas relaciones de fondo, se presenta realmente la cuestión de saber si ese concepto corresponde al nombre que se le da. Una secuencia de ácidos aminados puede tener un papel explicativo en genética, pero si la explicación genética exige que la secuencia sea, por ejemplo, una fuente de impulsos electroquímicos, la secuencia de ácidos utilizada en la explicación difícilmente pueda ser la de la teoría de fondo. En biología molecular nada da a secuencias de ácidos aminados contenidos en una molécula de ADN las propiedades de axones neuronales, de manera que las explicaciones genéticas deben respetar este hecho. Llamaré a los marcos teóricos de este tipo "redes en cadena" a causa de las maneras restrictivas en que están ligados sus conceptos centrales.

Prosiguiendo con el continuo, el punto siguiente está ilustrado por la teoría atómica de la materia. Se pueden tratar los conceptos que forman parte del fondo teórico en este caso de manera mucho menos restrictiva. Un átomo es la cantidad de materia más pequeña que junto con otras cantidades hace del

mundo macroscópico lo que nos parece ser. Por cierto, hay propiedades y relaciones que habitualmente se atribuían a los átomos: éstos han sido clasificados por tipos, son elementos que se mueven y sus movimientos están implicados en muchas de las apariencias específicas que el mundo nos presenta, etc. Esas propiedades bastan para restringir la gama de las teorías particulares que pudieran considerarse como partes del marco atomista. Pero dichas propiedades no son restrictivas hasta el punto de excluir de ese marco la concepción contemporánea de la materia. Hoy se atribuyen a los átomos propiedades que habrían dejado pasmado a Demócrito, pero éste no habría tenido más dificultades que nosotros mismos para reconocer el concepto de átomo de nuestra concepción contemporánea.

Considerando el carácter menos constrictivo del fondo teórico en este caso, se lo puede describir razonablemente como una red elástica de conceptos. Es posible extender la particular concepción que tenemos de un átomo mucho más de lo que podemos hacerlo en el caso de un ácido aminado, aunque por cierto un elástico también puede romperse. La teoría atomista del mundo habría podido conciliarse con la concepción humoral de las enfermedades; y habría podido revelarse tan errónea como las otras teorías que asignan propiedades suplementarias a los átomos y carecen de todo valor explicativo. Pero no fue esto lo que ocurrió.

El último ejemplo, el del lenguaje, es aquel en que los conceptos están ligados solamente de manera muy laxa con las teorías específicas de que dependen. A través de toda la historia de las teorías destinadas a explicar el lenguaje humano hubo numerosas formas de teorías gramaticales. Y por cierto sería difícil enumerar todas las teorías propuestas durante los treinta últimos años. Sin embargo en todas ellas los conceptos de frase, de nombre y de adjetivo han figurado de manera reconocible. El hecho de que esas diferentes teorías incompatibles entre sí utilicen un marco de fondo fijo de conceptos muestra hasta qué punto son poco estrictos los conceptos que forman parte del marco. Pues sería mucho más difícil especificar las propiedades que *deba tener algo* para ser una frase o un nombre en un lenguaje que especificar las propiedades que deba tener algo para ser un átomo (saber por qué esto es así sería un proyecto muy interesante, aunque ciertamente debería ser objeto de otro artículo).

Atendiendo a la nomenclatura que he utilizado hasta aquí, no debería sorprender que llame a este tercer tipo de fondo teórico una "red laxa". Las regularidades y propiedades que relacionan entre sí conceptos usuales, como los de frase, nombre y verbo, ejercen sólo una influencia muy débil en las teorías particulares de la gramática que utilizan esos conceptos para explicar el lenguaje humano. Y sin embargo, cualquiera que sea su carácter específico, cualquier gramática puede valerse de esos conceptos, con lo que se muestra hasta qué punto es mínimo el caudal teórico que tiene que ver con ella. El concepto de átomo puede reconocerse en las teorías de Demócrito, de Dalton y de la física contemporánea a pesar de las enormes diferencias internas de esas teorías. Pero hasta ese concepto es tan intrincado que es posible imaginar una teoría de la naturaleza que rechace el marco atomista. En cambio, considérese qué difícil es imaginar una teoría gramatical que, a causa de sus compromisos teóricos, nos lleve a pensar que no hay frases ni nombres ni verbos (no digo que no pueda hacerse, digo solamente que es más difícil en este caso que en el caso de la noción de átomo).

Después de estas observaciones algún tanto digresivas volvamos a los problemas principales de este párrafo. Por razones diferentes y con diferentes grados de entusiasmo, Stich y Churchland contemplan la posibilidad de que los desarrollos de la ciencia cognitiva puedan obligarnos a rechazar los conceptos de la psicología ordinaria. Según vimos, algunos de sus ataques, especialmente aquellos que consideran estancada la psicología ordinaria, sólo tienen sentido si se entiende por "psicología ordinaria" un marco teórico de fondo; no tiene sentido decir que teorías específicas de la psicología ordinaria están estancadas. Pero he tratado de mostrar que los marcos de fondo pueden tener relaciones muy diferentes con las teorías específicas que ellos autorizan. De conformidad con lo que dicen Stich y Churchland, muy bien puede ocurrir que conceptos de fondo, como creencias y deseos, puedan emplearse de maneras muy diferentes en psicología sin que esos empleos deformen la naturaleza intrínseca de dichos conceptos. En nuestro estado actual de saber, tal vez esos conceptos sean como el concepto original de átomo; hemos rechazado las teorías particulares en las que figura ese concepto sin rechazar el concepto mismo. (Dicho sea de paso: no puedo dejar de preguntarme si Stich y

Churchland desean que rechacemos el concepto de átomo porque éste ha hecho estancar nuestra teoría de fondo de la materia durante millares de años.) O, tal vez, los conceptos de creencia y de deseo son como los de frase, nombre y verbo: simplemente es difícil imaginar una teoría gramatical privada de tales conceptos, por grandes que sean las diferencias entre las particulares concepciones gramaticales. Y sin embargo, en otro sentido, tal vez Stich y Churchland tengan razón: quizá los conceptos de fondo se parezcan más a los conceptos de humores corporales de la antigua medicina.

Estas son cuestiones importantes y supongo que lo que he hecho sobre todo en este artículo fue formularlas sin tratar de responder a ellas, pues ésta sería tarea de otro artículo. Sin embargo vale la pena comenzar a responder aquí, lo cual nos ayudará a resolver mejor estas cuestiones ulteriormente y tal vez me permita influir en el lector aun cuando aquí no puedo ofrecer argumentos rigurosos contra las tesis de Stich, de Churchland (y de Dennett). Como soy incapaz de responder directamente a estas cuestiones, propongo abordarlas con algunas observaciones, primero sobre la psicología cognitiva, luego sobre la filosofía y por último sobre la psicología ordinaria.

## 5. La psicología cognitiva vista por un aficionado

Quisiera esbozar aquí lo que puede decir un aficionado de la investigación realizada en psicología cognitiva. Sin embargo mi finalidad es sencilla; quiero establecer qué papel desempeñan los conceptos de la psicología ordinaria en la psicología cognitiva. Expondré mi conclusión sólo en el último párrafo de este artículo.

Un buen punto de partida, aunque algún tanto caduco, es el conductismo. La verdad simple —si no ya simplista— es la de que la psicología cognitiva es la hija bastarda del conductismo. Como hija, debe su existencia a su padre, pero como hija problemática, sólo podía desarrollarse y prosperar una vez que su padre se hubiera retirado de la escena. No deseo detenerme aquí para considerar el primer punto, pero todo el mundo estará de acuerdo sobre el segundo. La psicología cognitiva floreció a partir del momento en que los psicólogos cognitivos ya no sufrían la presión de los métodos y de las hipótesis del conductismo. Una vez que se hubo aceptado plenamente que el

espíritu/cerebro era un objeto de estudio como las entradas y salidas energéticas de que aquél era considerado responsable, los hombres entregados a la teoría pudieron dar libre curso a su ingenio explicativo. Presuponiendo, tomando y aun arrebatando recursos conceptuales de otras esferas, construyeron modelos mucho más elaborados para forjar su concepción del espíritu. Así lo hicieron y continúan haciéndolo hoy, basándose en el hecho de que los métodos experimentales concebidos durante el pasado conductista se habían revelado tan frágiles como un simple castillo de naipes. Pues hasta el diagrama más complejo compuesto de casillas y de flechas que contenía nociones tomadas, ya de la filosofía, ya de la teoría de la computación y a veces de la psicología ordinaria, depende en última instancia de los descubrimientos experimentales. Se puede elaborar un modelo para la memoria partiendo de la idea de que hay algo común entre las bases de un ordenador y el espíritu. Pero al responder a todo un conjunto de hechos científicos sobre nuestras aptitudes de memoria, el modelo recibía un fundamento más sólido que el de la simple especulación.

Si consideramos diversas esferas de la psicología cognitiva, comprobaremos que muchas de las diferencias individuales que hay entre los distintos sectores de la indagación provienen de los tipos de conceptos utilizados en las estructuras teóricas, en las casillas y las flechas (cuando éstas son aplicables) de los esquemas explicativos particulares. No pretendo que este punto sea nuevo, pero quiero precisarlo prestando particular atención al uso de los conceptos de la psicología ordinaria en el esquema.

En la extremidad más vaga de la escala, tenemos los conceptos que no son explícitamente propios de la psicología ordinaria pero que parecen más próximos al idioma psicológico ordinario de lo que puedan hacer creer sus rótulos de corte técnico. Esto ocurre, por ejemplo, con las explicaciones del autismo que apelan a conceptos tales como los de *metarrepresentación*, y en el contexto en que se utiliza éste, el concepto se parece mucho a lo que todo el mundo llamaría una creencia sobre una creencia. Sin embargo aquí debemos estar atentos. El concepto de metarrepresentación puede ser en un contexto una forma de creencia, pero en otro contexto puede ser algo menos definido que una creencia. Por ejemplo, si vamos más lejos en este sentido, la noción de representación (la de una



frase ambigua, por ejemplo) tiene rasgos que parecen diferenciarla netamente de una simple creencia. En general (aunque no siempre) tenemos acceso consciente a nuestras creencias, en tanto que eso no parece cierto en el caso de las representaciones invocadas en lingüística.

Claro está, hay clases y clases de accesos. No tengo el tipo de acceso a representaciones verbales que me permitan hacer toda clase de inferencias sobre esas representaciones de la misma manera en que puedo hacerlo con mi creencia de que hoy es día viernes, pero por lo menos se puede sostener que tomo nota de ello. Se me podría mandar que dirigiera mi atención a esas representaciones por más que esto tuviera consecuencias desastrosas sobre mis facultades de conversación, pero eso sería posible. Sin embargo, a medida que uno progresa, aunque sin abandonar el dominio de la psicolingüística, la noción de representación se hace más esfumada: representaciones fonéticas de las palabras no son (cualquiera que sea el análisis que se haga de ellas) representaciones de las que yo pueda tomar nota y que pueda manipular. No puedo obedecer a la orden de dirigir mi atención a las representaciones fonéticas que penetran en mis oídos, salvo en el sentido no pertinente en que puedo oír lo que se me dice. Y lo que es válido en el caso de las representaciones fonéticas lo es aun más en el caso de las representaciones postuladas en las teorías del tratamiento de la información visual.

Resumamos; los conceptos con los que parece que tenemos que habérmolas en el dominio de la psicología cognitiva son conceptos teóricos cuyas afinidades con los conceptos de la psicología ordinaria son de grado variable. En la escala imaginaria que acabo de recorrer, el concepto de representación probablemente se encuentre a mitad de camino: parece situado por encima de los conceptos más descolantes de la psicología ordinaria, como los conceptos de creencia, y por debajo de conceptos de representaciones que figuran en los análisis de la información visual. Por lo demás, es perfectamente claro que los psicólogos cognitivos tienen muy poco tiempo para dedicar a cuestiones como la de la naturaleza de las nociones propias de la psicología ordinaria. Tal vez esto se deba a los objetivos explicativos de la psicología cognitiva y a sus orígenes conductistas.

## 6. Filosofía y psicología cognitiva

En mi examen de los tratamientos filosóficos de las nociones de la psicología ordinaria puedo avanzar con mayor rapidez sencillamente porque aquí me siento en un terreno más firme, ya que me encuentro en el seno de mi propio campo.

A pesar del breve episodio del llamado “conductismo filosófico”, probablemente sea justo afirmar que el conductismo nunca atrajo suficientemente a los filósofos como algo más que una curiosidad. Llevando al límite sus operaciones de rumiar sobre el problema cuerpo-espíritu, algunos filósofos perdieron el dominio sobre sí mismos hasta el punto de abrazar el conductismo por considerarlo una salida. Pero en general, para ser honesto, la principal atracción del conductismo para los filósofos de hoy está sobre todo en ejemplos útiles para mostrar en sus conferencias cómo no es posible resolver el problema cuerpo/espíritu. En realidad, a diferencia de los psicólogos, los filósofos nunca estuvieron realmente obsesionados por cuestiones prácticas referentes a las facultades y actividades de los seres humanos. Si se le preguntara a un filósofo ignorante por completo de la psicología cognitiva qué son las aptitudes humanas de razonar, de ver, de hablar, de navegar en espacios estrechos o de reconocer a su esposa, es muy probable que dicho filósofo diga que todas esas son aptitudes mentales, rasgos de la vida mental que posee toda *persona*, todo ser humano. Cuando el psicólogo se debate por hallar un esclarecimiento teórico del control que ejerce el espíritu/cerebro sobre nuestras aptitudes, los filósofos se contentan a menudo con pensar que es el yo o el espíritu lo que constituye la base de esas aptitudes; y aquí el yo se concibe como algo que tiene una integridad por encima de toda discusión.

Más que pretensiones explicativas, lo que sobre todo ha preocupado a los filósofos (mucho antes de que se interesaran en la psicología cognitiva) fueron cuestiones sobre la naturaleza de la creencia y particularmente sobre la manera en que ha de entenderse la función y la índole de frases que figuran en comunicaciones sobre contenidos de creencias. Verdad es que yo como filósofo me siento casi embarazado ante el número de brillantes espíritus que han dedicado su vida al intento de especificar cómo funciona una frase engastada dentro de una comunicación sobre la creencia. Y sin embargo hay muy poco en la bibliografía sobre cuestiones como: ¿Qué es una creencia?,

¿qué función tiene atribuir una creencia?, y ¿cómo funciona ésta en nuestro esquema de psicología ordinaria? Lo poco que hay en la bibliografía sobre este tema suele quedar relegado al margen. Quine especula que las atribuciones de creencias pueden entenderse así:

“Nosotros nos proyectamos en lo que nos imaginamos que fue el estado del espíritu del hablante (partiendo de sus observaciones y de otras indicaciones) y decimos lo que en el estado así simulado es natural y pertinente en nuestro lenguaje”. (Quine 1960, pág. 92)

Y Stich, entre otros, supone que en general esto es todo lo que hay que decir sobre el tema.

Muchos filósofos, en diferentes pasajes y de diferentes maneras, hacen referencia a las generalizaciones implícitas que gobiernan el concepto de creencia, pero dicen muy poco sobre el contenido de tales generalizaciones. Un caso particularmente claro es el de Fodor. A diferencia de Churchland, de Stich y de otros, Fodor es un ardiente defensor de la psicología ordinaria y sin embargo, en su reciente libro, *Psychosémantique*, ofrece un análisis meramente alusivo sobre el tema. En primer lugar, limita su campo de estudio a lo que llama las “actitudes”, luego las caracteriza por poseer los siguientes rasgos esenciales:

I) las actitudes son semánticamente evaluables

II) tienen fuerza causal

III) las generalizaciones implícitas de la psicología ordinaria, como deseo/creencia, se aplican ampliamente (pág. 10).

No es este el lugar apropiado para considerar los caracteres I y II, pero no deja de ser algún tanto decepcionante el hecho de que Fodor no diga sobre las generalizaciones implícitas de la psicología ordinaria nada más que esto: dichas generalizaciones incluyen los resultados de la teoría de la decisión y de las tesis como esta: “El gato está sobre la alfombra” es la forma habitualmente usada para comunicar que el gato está sobre la alfombra.

## 7. Dar su oportunidad a la psicología ordinaria

La conclusión de las dos secciones anteriores es la siguiente: los psicólogos cognitivos, en su intento de comprender el

desarrollo y el funcionamiento del espíritu/cerebro, de buena gana emplean conceptos específicos que evidentemente forman parte del marco de la psicología ordinaria. Al hacerlo, dan nuevos nombres a esos conceptos y los reelaboran a fin de conferirles nuevas propiedades y de alterar sus relaciones con otros conceptos explicativos. Pero en la psicología cognitiva en ninguna parte descubro un interés general por el esquema de la psicología ordinaria; no se encuentra en ella ningún intento de decir lo que hay de correcto o de incorrecto en el esquema total para explicar hechos, como por ejemplo el de que Barbara haya abandonado a Richard.

En cambio, los filósofos se ocupan realmente del marco en su conjunto, sólo que lo hacen de manera esquemática. Lo que interesa al filósofo son las propiedades lógicas de la representación en general y las maneras en que se pueda entender esa explicación atendiendo a representaciones. La creencia figura en las discusiones únicamente como un ejemplo y, que yo sepa, hay muy pocos trabajos filosóficos que puedan considerarse contribuciones a nuestra comprensión de la explicación detallada de la acción humana que da la psicología ordinaria. A decir verdad, no encuentro ningún pasaje en el que los filósofos se hayan planteado las cuestiones que me parecen esenciales respecto de la psicología ordinaria, como por ejemplo la cuestión de la importancia de nociones no evaluables semánticamente que entran en la comprensión de conceptos como el de creencia. De manera más específica, ¿cuál es la relación que hay entre conceptos, como rasgos de carácter, costumbres, necesidad e interés, y nuestros conceptos de actitudes proposicionales?

Lo que indica todo esto es que sencillamente no estamos en condiciones de decidir si la psicología ordinaria es un marco en cuyo interior pueda desarrollarse una comprensión futura de la cognición humana. Para servirme de mi terminología anterior, diré que hay por lo menos tres posibilidades sobre este particular:

I) Es una red en cadena que condiciona sus nociones centrales de manera tan rígida que si descubrimos que sus nociones violan las condiciones tenemos el derecho a dudar de que dichas nociones formen aún parte del marco inicial. Por ejemplo, si las condiciones que en psicología ordinaria pesan sobre el concepto de creencia y sus interacciones con los conceptos de deseo, de inten-

ción, etc. son suficientemente ricas, entonces será imposible considerar un estado cualquiera como una creencia si ese estado viola aquellas condiciones.

II) Es una red elástica que condiciona sus conceptos centrales de manera más o menos laxa. En ese caso, podríamos contemplar la posibilidad de un desarrollo conceptual que sea compatible con el marco inicial. Es decir, nuevas maneras explicativas de utilizar la noción de creencia sin que tengamos la impresión de que ya no estamos hablando de la noción de creencia.

III) Es una red laxa de conceptos que condiciona sus conceptos centrales de manera tan laxa que los conceptos utilizados en teorías explicativas podrían ir en toda clase de direcciones sin que esas teorías dejen de pertenecer al marco inicial. Hasta podría ocurrir que los nuevos desarrollos realizados con una terminología distintiva —una terminología que no corresponda precisamente a la psicología ordinaria tal como se la emplea corrientemente— no nos obligarán a abandonar la psicología ordinaria.

Me siento inclinado a considerar la psicología ordinaria como algo perteneciente a la tercera categoría, a considerarla como algo compuesto de nociones que (así como los conceptos de frase, de verbo y de nombre en lingüística) son virtualmente imposibles de rechazar y son sin embargo suficientemente flexibles para figurar en diferentes tipos de teorías explicativas. Desde luego, y como ya lo admití antes, no puedo probar aquí este punto, pero mis observaciones deberían por lo menos dar al lector razones para pensar que ese punto debe ser probado *antes* de que podamos decidir si una concepción del tipo de la de Stich o de Churchland es correcta. Los numerosos argumentos específicos que nos da Stich contra el concepto de creencia están casi exclusivamente fundados en observaciones bastante superficiales sobre la manera en que se utiliza ahora el concepto de creencia en contextos corrientes. El resorte de su argumento consiste en hacernos ver diversas clases de desfases entre la noción usual de creencia y las nociones de la psicología cognitiva. Pero para que los argumentos resulten aceptables debemos aceptar de antemano que la psicología ordinaria es como una red en cadena. Sólo con esta hipótesis tiene un sentido pensar

que el uso de la noción de creencia en las explicaciones corrientes no corresponde al uso de los conceptos aproximados en las ciencias cognitivas y que esto amenace a la psicología ordinaria. Todo lo que Stich nos dice es que bien pudiera ser que la psicología cognitiva, lejos de amenazar a la ordinaria, nos dé nuevas maneras de concebir lo que podríamos considerar felizmente como la psicología ordinaria futura.

Por supuesto, otra línea de argumentación que sigue Stich, y aun más vehementemente Churchland, consiste en afirmar que la psicología cognitiva misma se desarrollará de maneras tan ajenas a la psicología ordinaria que ya no habrá posibilidad de entenderlas relacionadas de las maneras antes sugeridas. Pienso aquí especialmente en la amenaza de las explicaciones que apelan a los contenidos, como las explicaciones de PDP (*Parallel distributed Processing*) o las explicaciones que apelan a las neurociencias. No puedo tratar aquí este punto pero quiero hacer notar brevemente dos cosas. En primer lugar, muchos defensores no filósofos de elementos tales como las redes PDP consideran que su trabajo recurre a la noción de contenido en la misma medida que los tipos de explicaciones más tradicionales de la psicología cognitiva. Y en segundo lugar, en la medida en que PDP y las redes neuronales se entienden como radicalmente diferentes de todo lo que se hizo con anterioridad, dichos autores tienden a atraerse las tajantes críticas de la misma comunidad de psicólogos cognitivos. Es cierto que en su forma cruda y filosóficamente ingenua las tesis de Churchland suscitan tanto aliados de la psicología cognitiva como defensores de la psicología ordinaria.

Las observaciones contenidas en este artículo, si es que muestran algo, se proponen alentar cierto grado de prudencia cuando se trata de tomar una decisión cualquiera en lo tocante al futuro de la psicología ordinaria. Pero antes de dejar al lector con esta simple advertencia, quisiera enumerar algunas consideraciones que, según me parece, favorecen la tesis de que el marco de la psicología ordinaria es suficientemente flexible como para permitir cambios de detalles sin cambios de sustancia.

I) Buen número de los trabajos actuales realizados en psicología cognitiva sólo tienen sentido si se los comprende como parte del marco de la psicología ordinaria

o como extensiones de ésta. Habría que analizar aquí cierto número de casos.

II) Aunque esto pueda discutirse, me parece razonable pensar que hemos aprendido una gran cantidad de cosas sobre los seres humanos con las teorías psicoanalíticas. Y la manera más plausible de entender dichas teorías es concebirlas como contribuciones a lo que ya hacemos en los ejemplos más elaborados de psicología ordinaria y concebirlas asimismo como desarrollos de lo que ya hacemos. Muchos ya lo han observado pero la significación que tiene esta observación sobre la flexibilidad de la psicología ordinaria no ha sido suficientemente apreciada.

III) La red de conceptos de la psicología ordinaria tiene una capacidad informal de cambio, capacidad que suele olvidarse con harta facilidad. Por ejemplo, una extensión sutil de nuestra manera de comprender a los seres humanos fue producida e incorporada en la psicología ordinaria por elementos tales como las metáforas derivadas de la jerga de los ordenadores. No es raro oír hablar de específicas conductas humanas que se atribuyen a “defectos” de constitución, concepto que tiene relación —aunque sea diferente— con la noción de hábito, propia de la psicología ordinaria y muy poco estudiada.

IV) Es muy difícil señalar lo que ha cambiado (o no ha cambiado) en la psicología ordinaria porque el marco explicativo pertinente a menudo no ha quedado registrado y porque las explicaciones particulares a menudo son efímeras o, cuando se han registrado, son elípticas. En tales circunstancias es fácil pensar que pocas cosas han cambiado. Los filósofos están demasiado inclinados a pensar que los griegos y los romanos se comprendían en los mismos términos que nosotros, pero esa suposición es totalmente injustificada. Una manera muy interesante de ver hasta qué punto se han producido cambios —y los filósofos habrían debido advertirlo— es la lectura de autores tales como Descartes, Spinoza y Hume sobre lo que consideraban pasiones humanas. Estos autores no están muy alejados de nosotros en el tiempo y vivían en el seno de instituciones sociales y

políticas más próximas a las nuestras que las instituciones, digamos, de los griegos. Y sin embargo, cuando uno lee textos como *Las pasiones del alma* de Descartes cobra la extraña sensación de estar en presencia de un ser humano diferente del que uno encuentra hoy. Pero una explicación diferente —la que yo estaría dispuesto a adoptar— es que la psicología ordinaria de Descartes no es la que emplearíamos hoy. Sin duda alguna Descartes explicaría la conducta del cartero en general de la misma manera que nosotros, pero ya he dicho que eso no prueba gran cosa.

### Referencias bibliográficas

- Churchland, P.M. 1981 "Eliminative materialism and propositional attitudes", *Journal of Philosophy*, LXXXVIII, 2, 67-90.
- Clark, A. 1986 «The Kludge in the machine», *Mind and Language*, III, 3, 45-63.
- Fodor, J. 1987 *Psychosemantics*, MIT Press.
- Hayes P. 1979 "The naïve physics manifesto", en D. Michie (comp.) *Expert systems in the micro-electronic age*, Edinburgh University Press.
- Quine, W.V.O., 1960 *Word and object*, MIT Press.
- Stich, S. 1983 *From folk psychology to cognitive science*, MIT Press.
- Wimmer H. y Penner J. 1983 "Beliefs about beliefs: representation and constraining function of beliefs in young children's understanding of deception", *Cognition*, 13, 103-128.

## Actitudes mentales y psicología del sentido común (contra la eliminación) \*

Radu J. Bogdan

### 1. Introducción

Excluyendo el empleo de la fuerza bruta, hay muchas maneras filosóficamente respetables de eliminar lo mental. Recientemente la estrategia de eliminación más popular se dirigió contra nuestra concepción común u ordinaria de lo mental. La estrategia se llamó materialismo eliminatorio (o posición eliminatoria). La motivación que está en la base de esta estrategia me parece ser la siguiente. Si la psicología del sentido común debe considerarse como la teoría que funda el dominio de lo mental, la teoría cuyo vocabulario y cuyos principios definen implícitamente lo que debe considerarse mental, entonces eliminar la teoría equivale a eliminar aquello de que ella habla. Si se demuestra que la teoría es falsa, luego aquello de que habla no existe. En otras palabras, si se puede mostrar que la psicología del sentido común no describe ni explica nada de la cognición humana, lo mental mismo es una ficción.

Al comienzo, esta postura representada por Rorty y Feyerabend entre otros se concentró en los conceptos de cognición del sentido común y en los aspectos sensoriales y fenoménicos de este concepto. Desde un punto de vista táctico, esto tenía sentido porque es en la arena fenomenológicamente accesible de la conciencia donde durante siglos colocó el espíritu el

\* De próxima aparición en *Noûs*. En la edición original francesa este artículo fue traducido del inglés por P. Engel, con autorización del profesor H. N. Castañeda y del autor.

paradigma cartesiano dominante. En los años recientes, sin embargo, el peso crítico de la posición eliminatoria se dirigió contra la concepción del mundo del sentido común y contra estados informacionales y funcionales como la creencia, el pensamiento, el deseo, etc. Esta redistribución del acento crítico tiene igualmente un sentido desde el punto de vista eliminatorio, porque es en el dominio del tratamiento cognitivo de la información y de las funciones cognitivas donde se concentra la investigación científica en este momento. Como ahora son las actitudes proposicionales —como las entiende la psicología del sentido común— las que definen implícitamente lo mental (en el sentido no fenoménico), la estrategia de la posición eliminatoria consiste en crear un conflicto entre las ciencias de la cognición y la psicología del sentido común, conflicto que, atendiendo a sus respectivos objetivos, se convierte en un conflicto entre cognición y espíritu o entre lo cognitivo y lo mental.

Por supuesto, es crucial la cuestión de saber si la postura eliminatoria tiene razón tocante a la psicología del sentido común y de su objeto, lo mental. A través de toda su historia, lo mental reapareció una y otra vez con un nuevo ropaje después de haber estado expuesto y luego eliminado con un ropaje antiguo. La razón de ello es que lo mental tiene funciones firmes, quizá permanentes, en el seno de nuestro esquema de comprensión corriente. Lo mental me distingue y me protege (como a todo individuo) del mundo circundante (subjetivo/objetivo) de mí mismo (¿de alguien? ¿de otro?); de la animalidad y de las máquinas, de la determinación física y social (libertad, creatividad, personalidad, individualidad); del simple tratamiento de la información y de la computación (ser humano/ordenador). Enumero todos estos hechos familiares para recordar lo que entra en juego en el debate sobre la psicología del sentido común y para sugerir la motivación (a veces oculta) que tienen muchos filósofos cuando quieren acudir en auxilio de la psicología del sentido común, la que estiman como la última barrera de defensa contra el avance ineluctable de la ciencia, del materialismo y de la cognición mecánica.

Sin embargo, no me ocuparé de la oposición que hay entre postura eliminatoria y psicología del sentido común desde ese punto de vista. Si bien estoy programado, como todos nosotros, para ser un psicólogo del sentido común, no soy un abogado de

su sabiduría ni un defensor de sus conceptos favoritos. Además, contrariamente a lo que piensan muchos filósofos, ser un psicólogo del sentido común no hace de nosotros expertos en psicología del sentido común. De manera que no tengo ninguna teoría de la psicología del sentido común; ésta sería ulteriormente asunto de la psicología social y de la antropología, según supongo. Con todo, tengo tres hipótesis de trabajo, una sobre la psicología del sentido común, otra sobre lo que falla en la posición eliminatoria y una tercera sobre lo mental. Quisiera reunir estas tres hipótesis; la imagen de conjunto resultante es en general la siguiente.

La posición eliminatoria da una falsa representación de la psicología del sentido común porque ofrece una falsa representación de la cognición misma. En particular, dicha postura da una falsa representación de las nociones corrientes de actitudes proposicionales y en consecuencia de la noción de lo mental que aquéllas definen implícitamente, porque representa falsamente las formas en las que la información es codificada, transformada en actitudes y tratada en la cognición central. Esas falsas representaciones se ven vigorosamente alentadas por el axioma de que todo lo que tiene que ver con la cognición puede explicarse de manera científica. Hay razones para dudar de la verdad de este axioma. Estas razones están objetivamente motivadas por los modos en que la información es ordenada y manipulada en la cognición central. Llamaré *mental* a esta forma de información que es específica de la cognición central. Se puede mostrar que los conceptos comunes de actitudes proposicionales (cuando se las construye correctamente como actitudes relativas a una información mental) representan verdaderamente algo importante sobre la cognición central. En suma, la argumentación contra la posición eliminatoria y en favor de la psicología del sentido común se apoyará en un argumento que nos lleva *desde* una nueva caracterización de la información mental y de las actitudes *a* la incapacidad de las formas corrientes del análisis científico para dar cuenta de ellas y luego *a* la idea de la legitimidad descriptiva del esquema que nos ayuda a localizar y a comprender la información mental y las actitudes, es decir, el esquema de la psicología del sentido común. La médula de la argumentación es la de que si la cognición central funciona de conformidad con la manera en que los conceptos de información mental y actitudes mentales suponen que ella

funciona, luego no pueden eliminarse esas nociones. Y si la psicología del sentido común es la fuente bien fundada de nuestros conceptos de información mental y de actitudes, entonces deben escapar a la eliminación.

Esta conclusión parece correcta, no porque yo considere que la teoría del sentido común es la teoría válida de la cognición central ni porque considere que sus conceptos de información mental y de actitudes son las únicas nociones correctas que se refieren a la cognición central. Pues en verdad no supongo nada de todo eso. No trato la psicología del sentido común como una teoría, sino que la considero como una práctica. En consecuencia, no considero que sus conceptos constitutivos describen correctamente lo que ellos caracterizan implícitamente. Los considero correctos prácticamente. El hecho de ser prácticamente correctos en cuanto a la naturaleza del espíritu no priva a los conceptos del sentido común de su propiedad descriptiva. Una práctica cualquiera alcanza sistemáticamente éxito sólo si las nociones que la guían representan las propiedades reales de los objetivos de esa práctica. Por consiguiente, el sentido común no es ni una teoría falsa de la cognición cuando se lo considera desde el punto de vista de su propiedad descriptiva, ni una simple heurística, cuando se lo considera desde el punto de vista de su propiedad instrumental, según ciertas versiones de la posición eliminatoria.

Y eso es lo que quiero sostener en las páginas que siguen.

La argumentación será relativamente compleja. Trazar sus líneas generales nos ayudará a seguir sus diferentes etapas y a ver como éstas sirven al objetivo de conjunto. En primer lugar identifico la índole del programa eliminatorio y distingo en él varias estrategias eliminatorias importantes. Esas estrategias son también análisis importantes de la psicología del sentido común. Este tema abarca la sección 2. La sección 3 examina el modelo de la cognición (encarado por todas las estrategias eliminatorias) que satisface los requisitos conceptuales y explicativos de las ciencias habituales de la cognición. El modelo, el de la cognición transductiva y modular (en adelante, transmodular) es el modelo que la posición eliminatoria opone al concepto de espíritu, propio del sentido común. La sección 4 examina de nuevo la noción de información en la que reposa todo lo demás, y distingue una nueva forma, la información mental, de la que se muestra que es psicológicamente real,

causalmente eficaz y constitutiva de diversas actitudes mentales. También muestro que la información mental es una forma de información irreductible en relación con las leyes transmodulares de la cognición y en consecuencia que no está determinada por esas leyes y que por lo tanto no puede explicarse por las ciencias de la cognición transmodular. La sección final, 5, vuelve a colocar la psicología del sentido común en el centro del cuadro. Se sugiere aquí que hay un análisis de la psicología ordinaria, asimismo emprendido por ciertos eliminadores; me refiero al análisis individualista, que se aproxima mucho a una concepción de la psicología del sentido común entendida como práctica y motivada por consideraciones biosociológicas, práctica que manipula la información y sus actitudes asociadas. Contrariamente al análisis eliminatorio, considero que esta práctica describe algo importante sobre el espíritu humano. Frente a la posición eliminatoria, lo mental vuelve a ocupar de nuevo el primer plano y el sentido común de nuevo queda en una posición legítima, la de una práctica y no la de una teoría.

## 2. Un diagnóstico de la posición eliminatoria

1. *La argumentación.* La argumentación eliminatoria nos da la fórmula básica siguiente para eliminar lo mental:

1) HIPOTESIS: la cognición es información calculada, computada y representada.<sup>1</sup> La representación y la computación de la información son rasgos objetivos sometidos a leyes, seres dotados de cognición en interacción con su ambiente, lo cual explica por qué esos rasgos pueden y deben ser caracterizados y explicados en una forma apropiada por las ciencias legítimas de la cognición.

2) HIPOTESIS: el sentido común caracteriza lo *mental* manifestado en estados cognitivos tales como creencias, pensamientos o deseos. Los filósofos consideran esos estados como actitudes relativas a contenidos semánticos y los llaman *actitudes proposicionales*. Los contenidos de actitudes y sus leyes constituyen la ontología bien fundada de especies mentales y de leyes que la psicología del sentido común utiliza para conceptualizar y explicar la cognición y la conducta.

3) HIPOTESIS: únicamente lo que puede explicarse *científicamente*, atendiendo a especies y a leyes empíricamente determinadas, es *real* en un nivel apropiado de abstracción. Todo lo demás es epifenoménico.

4) HAY QUE MOSTRAR: que lo mental no tiene *nada* que ver con lo cognitivo. Hay que mostrarlo sosteniendo que las nociones ordinarias capitales de actitudes proposicionales, que definen implícitamente lo mental, no describen nada *real* de la cognición o sobre la cognición.

5) HAY QUE MOSTRAR TAMBIEN: que las ciencias de la cognición pueden explicar *todo* lo que es importante sobre la cognición sin emplear nunca nociones del sentido común tales como las de actitudes proposicionales.

6) CONCLUSION: la psicología del sentido común no identifica ni explica nada real sobre la cognición. Lo mental (aquello que la psicología del sentido común define implícitamente) no es científicamente real. Lo cognitivo (aquello que las ciencias de la cognición definen implícita o explícitamente) excluye lo mental. Por consiguiente lo mental no existe. La eliminación queda cumplida.

2. *Versiones de la eliminación.* La argumentación eliminatoria presenta varias versiones distintas en cuanto a la manera en que la ciencia elimina el sentido común al purgar lo cognitivo de lo mental. Distingo tres versiones, formuladas en los términos en que ellas enfocan la psicología del sentido común y sus construcciones básicas. Identifico y discuto esas versiones sólo en relación con algunos parámetros que considero pertinentes para la argumentación. No me propongo entrar aquí en los detalles ni ser exhaustivo ni exacto. Resumo los aspectos centrales en un cuadro [véase el cuadro de la pág. 78] que basta por sí mismo y agrego algunas explicaciones aquí y allá. La versión *realista* puede resumirse del mejor modo con las palabras de su autor más descollante, Paul Churchland. Según esta concepción, la eliminación (realista) es:

“la tesis de que nuestra concepción común de los fenómenos psicológicos constituye una teoría radicalmente falsa, una teoría tan fundamentalmente defectuosa que los principios y la ontología de dicha teoría terminarán por ser desplazados, antes que progresivamente reducidos, por la neurociencia llegada a la madurez.”<sup>2</sup>

versión parámetros	realista	interpretacional	doble aspecto	individualista
forma	teoría empírica	teoría interpretativa	teoría empírica + evaluación semántica	historias individuales partiendo de un esquematismo
especie de base = actitudes proposicionales	procesos y estados internos	aspectos normativos y heurísticos	estados internos + sus referencias	admisión ad hoc de estados y aspectos internos
	actitudes frente a versiones intencionales / semánticas (proposiciones)			
leyes básicas	leyes empíricas	condiciones racionales / normativas	leyes empíricas	no hay leyes explícitamente formuladas
	leyes de contenido (causales, lógicas racionales, normativas)			
explicación	deductiva nomológica/funcional	intencional	cualquiera	comprensión individualizada
	científica: de leyes a <i>explananda</i> genéricas			
sujeto básico	lo mental = contenidos intencionales / semánticos			lo mental ¿cómo?
la psicología ordinaria es	protociencia			protoarte
eliminarios	Churchland	Davidson; Dennett	Putnam, Stich	Morton, Wilkes

Esto implica que “la neurociencia llegada a la madurez” explicará todo lo que hace falta saber de la cognición humana y en ese caso la psicología del sentido común reflejará que ha sido tan sólo un dispositivo intelectual cómodo, transitorio pero estéril, el subproducto pasajero de nuestra ignorancia.

Atendiendo a los parámetros elegidos para mostrar cómo la posición eliminatoria considera el sentido común, puede reconstruirse del modo siguiente el argumento que está en la base de la versión realista. La psicología del sentido común es una *teoría empírica* que describe la cognición humana, de ahí su sentido realista. Tiene su propia ontología de especies y de leyes que, como toda teoría que se respete, suministra sus recursos explicativos principales. Se supone que las *especies* (las actitudes proposicionales) caracterizan tipos fundados de estados internos tales como las creencias o los pensamientos. La psico-

logía del sentido común contiene una rica red de *leyes* sobre actitudes proposicionales, que no son ni analíticas ni normativas. Esas leyes describen regularidades objetivas en las actitudes o entre las salidas (*outputs*) verbales o de conducta, fundadas en relaciones lógicas entre los *contenidos* de las actitudes en cuestión. Las *explicaciones* del sentido común son del tipo deductivo nomológico, pues someten las actitudes particulares a las leyes correspondientes. Considerando todas estas cosas, puede estimarse que la psicología del sentido común entra en *competencia* con las ciencias fundamentales de la cognición (en particular las neurociencias), pues trata de explicar e identificar los hechos cognitivos y de conducta. El resultado de la competencia es claro. La psicología del sentido común es una teoría irremediabilmente inadecuada, superada en su propio dominio por las neurociencias. Por consiguiente, las neurociencias *desplazarán* y *reemplazarán* sus nociones explicativas. El espíritu desaparece y la cognición ocupa su lugar.

La versión *interpretacional* considera la psicología del sentido común también como una *teoría*, concebida, ya como una anticipación (Dennett), ya en continuidad con ciertas ciencias sociales/normativas apropiadas (Davidson), pero no como una teoría empírica. Su tarea es interpretar, racionalizar y evaluar la cognición a la luz de valores tales como los de optimalidad, utilidad o racionalidad. Como teoría, la psicología del sentido común está provista de su propia ontología de especies y de leyes, pero no se supone que las *especies* (las actitudes proposicionales que representan los aspectos normativos o teleológicos de la cognición) ni las *leyes* (que incorporan condiciones normativas o teleológicas en las relaciones que hay entre los contenidos) designen tipos reales ejemplificables de rasgos cognitivos o de regularidades. Cualquiera que sea su estructura lógica particular, una *explicación* será intencional, en primer lugar porque trata lo mental como algo opaco, tal como está representado por el agente, y en segundo lugar porque la explicación está concebida para hacer aparecer el *explanandum* como razonable u óptimo, según el saber y los intereses del agente y las normas sociales a las que éste se atiene. De manera que no puede haber competencia entre las ciencias empíricas de la cognición y la psicología interpretativa del sentido común porque no comparten la misma ontología de especies y de leyes. Las actitudes proposicionales y sus leyes no



tienen realidad cognitiva. En esa medida quedan *eliminadas* del trabajo de la cognición, por más que continúen siendo importantes desde el punto de vista de la explicación y de la predicción. El espíritu, tal como está implícitamente definido por el sentido común, es una construcción explicativa heurística aunque no designa nada. El concepto de espíritu, por más que no forma parte de la cognición y en consecuencia sea irreal, nos ayuda a comprender y a predecir la cognición.

La versión del *doble aspecto* nos cuenta una historia más amplia sobre lo que hace la psicología del sentido común. El aspecto interno controla lo que pasa cognitivamente en la cabeza, en la forma típica de papeles conceptuales representados por relaciones lógicas entre los contenidos. Se supone que este aspecto es insuficiente para fijar los contenidos semánticos de las actitudes. Un aspecto externo, que controle las relaciones causales con el ambiente, es igualmente necesario para establecer el carácter referencial semántico (*semantic aboutness*) de los contenidos. La psicología del sentido común de doble aspecto conecta los estados mentales con sus blancos semánticos reales, ya sea en términos de referencias, ya sea en términos de condiciones de verdad o de convenciones sociales. Los conceptos de actitudes proposicionales representan juntamente las *especies* de estados internos y sus *coordinadas* semánticas externas. Cualquiera que sea la naturaleza de las leyes y en consecuencia de las explicaciones del sentido común, ellas no se refieren ciertamente a la cognición como tal. La noción de lo mental es un concepto híbrido, explicativo/evaluativo, que en la cognición no corresponde a ninguna articulación natural de ésta. Aquí la eliminación de lo mental sigue una reconceptualización estipulada del conjunto de las actitudes y de los contenidos semánticos.

Por fin, la estrategia *individualista* ataca en otra dirección. Esta estrategia tiene menos motivaciones teóricas que las otras y sin embargo aflora aquí y allá en la bibliografía.<sup>3</sup> Trata el sentido común como una práctica social que nos suministra instrumentos con los que fabricamos estudios de casos particulares sobre la cognición y la conducta de un agente. A diferencia de las otras tres estrategias de eliminación, que presentan a la psicología del sentido común como una especie de protoarte o de aptitud que suministra recetas imaginativas para tratar de manera única lo que es individual o nuevo o diferente. Así

concebida, la psicología del sentido común *no es una teoría* en ningún sentido. Sus conceptos no representan *ninguna especie* empírica o normativa, es decir, ninguna especie. El concepto de especie abarca una clase heterogénea cualquiera con alguna propiedad distinta. Las actitudes proposicionales interpretadas de manera individualista no son de esta naturaleza. Su esencia está definida implícitamente por nuestras atribuciones y evaluaciones ordinarias en contextos apropiados. Si consideramos atentamente esas atribuciones y evaluaciones, lo que comprobamos es una mezcla pragmática de dimensiones desunidas, y a veces sin relación unas con otras, que sirven para fijar los contenidos y las actitudes en cuestión. Como ninguna naturaleza única se puede discernir en esos contenidos y actitudes, no se ve muy bien lo que sean de un contexto a otro y eso no parece tener importancia. *¿Leyes?* Por supuesto, dirá el individualista, el sentido común se vale de todas las generalizaciones que puedan servir a sus objetivos. Pero esto no quiere decir que el sentido común formule explícitamente su *propio* conjunto de leyes. Puesto que la noción de contenido es fluida, sería muy difícil tener verdaderas leyes de contenido de todas maneras. *¿La explicación?* Por supuesto, hay explicación de cierta clase, pero no se manifiesta aquí ningún criterio lógico particular de explicación. Todo cuanto importa es aplicar hábilmente bastantes detalles pertinentes hasta que un perfil individual de un elemento particular de nuestra cognición o motivación o conducta se manifieste y tenga sentido en relación con las normas que prevalecen en el fondo cultural. Evidentemente aquí no hay una competencia con la concepción científica de la cognición. De cualquier manera, lo mental queda *eliminado* pues ahora está más alejado que nunca de lo cognitivo con el cual ya no puede comparárselo.

### 3. El perfil eliminatorio de la cognición: continuación del diagnóstico

El partidario del programa eliminatorio, para defender su posición, debe no sólo presentar lo mental (en la forma de las actitudes proposicionales) como epifenoménico y lo cognitivo como real porque el segundo y no el primero es científicamente caracterizable y explicable, sino que *también* debe mostrar que ningún aspecto explicativo importante de la cognición está, en

principio, fuera del alcance de las ciencias de la cognición, pues en el caso de que hubiera tales aspectos, éstos podrían legitimar en principio modos de teorización y de explicación diferentes de aquellos que promueve la ciencia. De modo que la posición eliminatoria debe defender la idea de que la cognición es un objeto susceptible de ser caracterizado y explicado científicamente (más exactamente por las ciencias cognitivas) *en su totalidad*, sin residuo alguno. Y sin embargo los requisitos habituales que pesan sobre la explicación en las ciencias de la cognición ponen serias limitaciones en el perfil de la cognición favorecida por esta postura. ¿Puede el eliminador jugar en dos campos? ¿Puede tratarse científicamente toda la cognición? Esa es la pregunta a la que hemos de responder en esta sección.

### 1. Modelación de la explicación: clasificar, establecer leyes y explicar

Dar una caracterización científica de un dominio de hechos es un proceso de dos etapas. La primera es la de la conceptualización. Consiste en organizar los hechos en especies (o tipos) en cuyo seno pueden establecerse leyes persistentes de correlación. La primera etapa es un preludeo a la segunda, la de la explicación, en la cual se localizan los hechos particulares que queremos comprender dentro de la ontología elaborada en la primera etapa. Así caracterizamos (o clasificamos en tipos) esos hechos según lo que indica la ontología, y luego los sometemos a leyes apropiadas. Cuando esta operación lógica queda terminada hemos explicado los hechos. Y si todo esto se hace según la vía científica, hemos explicado científicamente los hechos.

Para ser un objeto de explicación científica, la cognición debe tratarse de manera que pueda aplicarse a ella esta clase de explicaciones, lo cual quiere decir que la explicación debe pasar el test científico: debe satisfacer ciertos requisitos ordinarios de la explicación científica. Aquí no intentaré hacer ningún análisis exhaustivo de esos requisitos. Todo aquello de que tenemos necesidad en la presente argumentación es tan sólo tener en cuenta los requisitos que revelan la índole del enfoque científico de la cognición y las limitaciones de éste.

La explicación científica es una explicación que *parte* de especies y de leyes: los hechos que han de explicarse deben tener una relación lógica con los hechos y las leyes invocados, es decir,

una relación que satisfaga nuestro sentido y nuestro deseo de explicación. La tradición quiere que la forma más respetable de explicación científica sea *deductiva nomológica*. Esta forma clasifica el explanandum como una instancia particular de una especie y (una vez especificadas las condiciones contextuales) la deduce de una ley. Esta forma parece dar mejores resultados en los niveles ontológicos inferiores, por ejemplo físicos y químicos; en los niveles más complejos, donde las estructuras implican facultades y disposiciones, tenemos necesidad de una forma de explicación *morfológica* con la cual explicamos las facultades atendiendo a estructuras subyacentes y aptitudes fundamentales. El tipo morfológico parece el más apropiado para casos complejos de composición física (*hardware*), incluso hasta la base neuronal. Sin embargo, cuando las estructuras y las facultades están destinadas a cooperar para producir un resultado o realizar una función, tenemos necesidad de una nueva forma de explicación, de tipo *sistemático*.<sup>4</sup> Lo cual no es sorprendente, pues ése es el nivel de explicación requerido para los fenómenos cognitivos.

Cualquiera que sea su forma, la explicación científica sólo da resultado si su dominio está concebido (en un nivel apropiado de abstracción) de manera tal que permita una relación lógica entre el explanans y el explanandum. Esta condición se satisface, según vimos, si la *clasificación y las leyes* de la ontología apropiada nos aseguran que los hechos particulares del dominio en cuestión corresponden a especies cuyas propiedades y relaciones están regidas por leyes referentes a correlaciones o a facultades primitivas o a estructuras o a funciones, según los casos. Mostrar que el explanandum es un caso particular de una especie, sometido a una ley cualquiera, nos asegura que aquello que sabemos de la especie y de la ley se extiende a ese caso. Esto es lo que ocurre en toda explicación científica.

Otro requisito es el del *materialismo*. Trátase de un requisito metafísico que prescribe que todas las entidades particulares (*tokens*) sean materiales (cosas del espacio-tiempo) en un nivel cualquiera de abstracción y que todas las especies se reduzcan a una entidad material en un nivel inferior cualquiera. Todo es composición física. En nuestro mundo, el fisicismo es la versión principal del materialismo: todo es de composición física.

El requisito siguiente, que yo llamo el requisito del *explanandum genérico*, es a la vez delicado y muy central en mi argumento.<sup>5</sup> Este requisito postula que en la explicación científica de un hecho *individual*, cosa o propiedad, el *explanandum* es un individuo *genérico*. Veamos lo que quiero decir con esto. Toda entidad material particular es *trivialmente* única porque el conjunto de *todas* sus propiedades la hace diferente de cualquier otra entidad. Si se consideran todas las propiedades de una entidad particular, luego dos entidades particulares no son nunca semejantes, sino que cada una es (trivialmente) única. Pero la explicación científica no se preocupa por todas las propiedades de una entidad particular. Y esta indiferencia es deliberada. La unicidad trivial no puede ser explicada ni tiene necesidad de serlo. Al buscar lo universal en lo particular, la explicación científica hace abstracción *precisamente* de las propiedades que hacen única a una entidad particular.

Toda entidad particular que deba explicarse puede ser a la vez única y genérica. Es trivialmente única cuando se tienen en cuenta todas sus propiedades; es genérica cuando se ignora aquello que la hace trivialmente única. Una entidad particular es genérica cuando, en el caso de un contexto dado de causalidad temporoespacial, toda entidad de la misma especie, que ocupa la misma posición de *explanandum*, recibe la misma explicación. En cuanto a la explicación científica, lo realmente único es el contexto, no aquello que lo ocupa. El contexto identifica a su ocupante para *explanans* y en este sentido le confiere su unicidad trivial. El contexto mismo no está explicado. Lo explicado es el ocupante, pero sólo como entidad genérica, como representante anónimo de su especie. Dentro del marco de la explicación científica, toda partícula, molécula o cristal de tamaño idéntico o de constitución idéntica se comportará de la misma manera que cualquier otro sin importar el contexto temporoespacial que ocupe. Se trata de individuos genéricos porque, según la teoría, son miembros anónimos, sin originalidad, miembros medios y regulares, de especies uniformes, homogéneas y gobernadas por leyes. Como se ve, la cuestión no se refiere a lo que hay en la realidad, sino que se refiere a la manera en que las teorías modelan lo que hay en la realidad para explicarlo.

El concepto de *explanandum genérico* debe luego distinguirse del concepto de un *explanandum* no genérico o *esen-*

*cialmente único*. Esta última noción se refiere a objetos de explicación que son esencialmente únicos porque los contextos que ocupan intervienen de manera crucial o porque los *explananda* mismos tienen propiedades y disposiciones que son responsables de un comportamiento no sometido a leyes, un comportamiento creador y original. En general ambos van juntos: los contextos tienen una importancia explicativa cuando sus ocupantes muestran reacciones a estímulos causales, reacciones originales y no sujetas a leyes. Esto es lo que distingue la unicidad trivial de la unicidad esencial; los contextos temporoespaciales confieren una unicidad trivial a un *explanandum* dado cuando su comportamiento continúa siendo invariante en cambios de contextos; cuando el comportamiento no es así invariante, es sensible al contexto; si es esto lo que cuenta en la explicación, tenemos entonces un *explanandum* esencialmente único.

## 2. La explicación teórica cognitiva

La explicación teórica cognitiva es una clase de explicación científica. Es la explicación científica de la cognición en niveles apropiados de abstracción ontológica, según formas apropiadas a esos niveles. Los niveles mismos, tales como los definen las ciencias de la cognición (la psicofísica, las neurociencias o la psicología cognitiva) especifican los dominios de la explicación cognitiva teórica. Además de los requisitos generales tratados en la sección anterior sobre la explicación científica, la explicación teórica cognitiva debe satisfacer también cierto número de requisitos específicos que la hacen *cognitiva teórica*.

Un primer requisito afecta la construcción misma de la ontología de la que tiene necesidad la explicación teórica cognitiva. Ese requisito tiene que ver específicamente con las especies cognitivas que habrán de figurar en las conceptualizaciones y en las explicaciones de las ciencias de la cognición. Se conoce este requisito como *solipsismo metodológico* y estipula que clasificamos las especies cognitivas tales como las representaciones, las creencias o las intenciones sin tener en cuenta aspectos reales del ambiente a los que esas especies (en cada caso) pudieran hacer referencia. Esto no excluye la posibilidad de que se tengan en cuenta influencias sistemáticas del ambiente, de la sociedad o de la evolución (con frecuencia se cree erróneamente que eso es lo que implica este requisito). Es

menester tener en cuenta esas influencias si se quiere comprender lo que la cognición realiza en este mundo y cómo lo hace. El requisito excluye solamente las coordenadas *presentes y particulares* de la clasificación de las especies cognitivas.

También está el requisito del *internalismo*, el cual estipula que toda transacción informacional se realiza de manera interna en virtud del estado del organismo. Aquí la motivación es simple. Explicar hechos cognitivos y de conducta es explicar lo que rige causalmente la cognición y la conducta. La información es esa fuerza rectora. Una información particular es causalmente eficaz en un organismo si está convenientemente codificada por los estados del organismo.

### 3. ¿Qué es lo que queda explicado y hasta qué punto?

Reunamos ahora estos requisitos y veamos qué clase y qué cantidad de cognición permiten explicar de manera cognitiva teórica. Esto nos dirá a su vez hasta qué punto nuestra comprensión del espíritu cognitivo encuentra en definitiva su fuente en las ciencias de la cognición. Formularé la cuestión de la manera siguiente: ¿cuál sería la cognición humana si tuviera que satisfacer las condiciones de la explicación cognitiva teórica? ¿O podría encontrar la explicación sus especies y sus leyes para aplicarlas a sus explananda?

Comencemos por considerar la segunda pregunta. La respuesta debe hacer referencia a la composición física de un organismo y a su intención semántica principal: estos elementos solamente nos dan las especies y leyes requeridas por las ciencias de la cognición en diferentes niveles. La intención semántica básica de la cognición está esencialmente incorporada en transductores, en módulos que engendran las representaciones y en los centros motores y de control. Son los componentes que permiten al organismo organizar y tratar las informaciones de entrada al dirigir su conducta hacia aspectos pertinentes del ambiente. Puede decirse entonces que la información recibida por el organismo está de conformidad con sus aspectos e intereses y por consiguiente que es semántica. Por falta de espacio en este artículo debo ser dogmáticamente breve en este punto. Se encontrarán detalles en (3).

Pasemos ahora a la primera pregunta. Si la cognición humana fuera la cognición transmodular (transmitida por componentes materiales apropiados) y fuera suficiente para

poner en marcha los centros motores y de control y para guiar la conducta a diversos tipos de blancos, la explicación cognitiva teórica se aplicaría plenamente y con éxito a *cualquier* cognición particular y a *cualquier* conducta particular. Si la cognición fuera sólo cognición *semántica* (en el sentido naturalizado expuesto en un párrafo anterior), las ciencias de la cognición bastarían por sí solas para explicarla. También se puede formular este punto de otra manera. Supongamos que todo cuanto importa (cuando se trata de comprender la cognición) sea la manera en que la información es convertida (o transducida) y codificada (o representada) a fin de suministrar un mapa semántico de la conducta, en el sentido en que es la información lo que dice a los centros motores y de control qué clases de blancos deben apuntarse o deben hacerse reaccionar. La información que hace este trabajo hace un trabajo semántico. Además es el género de trabajo (así lo sugiero) que las ciencias de la cognición son perfectamente capaces de conceptualizar y de explicar. Lo que quiero explicar es que la cognición transmodular es esencialmente una cognición semántica, pues entrega, a través de regularidades legales, representaciones semánticas del ambiente y, por lo tanto, las ciencias de la cognición pueden comprenderla y explicarla como tal.

¿Por qué la cognición semántica es un objeto apropiado de la explicación cognitiva teórica? Porque toda instancia particular de ella puede ser explicada así. Pero ¿por qué esto es así? Porque toda información semántica particular (cuando satisface los requisitos de la explicación cognitiva teórica) corresponde a clases de componentes materiales y/o de intenciones semánticas que la hacen genérica y legal en posición de explanandum. Todos nosotros somos semejantes en nuestra cognición transmodular. Así estamos constituidos. Si encontramos las especies y leyes convenientes que se encuentran detrás de la manera en que estamos constituidos, nos vemos envueltos en la explicación cognitiva teórica. Y esto es precisamente lo que han hecho los hombres de ciencia durante los últimos decenios en los dominios de la percepción, del lenguaje, del aprendizaje, de la memoria, etc. Lo que tienen en común los esfuerzos de esos hombres es una comprensión de las maneras en que la información es transducida, codificada y manipulada y puesta así al servicio de los objetivos semánticos del organismo. (Es decir al

servicio de aquello a lo que debe ser constitutivamente sensible la conducta.) La concepción de Fodor (10) de la cognición transmodular explica este punto mejor de lo que yo puedo hacerlo aquí.

#### 4. *Recapitulemos*

Si la cognición transmodular es el único objeto de la explicación cognitiva teórica, entonces, para el partidario de la eliminación, debe ser la teoría de la cognición transmodular lo que define (implícitamente) lo *cognitivo* y excluye lo *mental*. Lo transmodular no es lo mental o, para decirlo de otra manera, el espíritu no puede ser transmodular. Esta es una conclusión que tiene sentido. Hasta aquí la posición eliminatoria tiene razón. Pero recordemos la otra parte de su proyecto: mostrar que la *totalidad* de la cognición puede explicarse de manera cognitiva teórica. Y sin embargo lo que puede explicarse así es tan sólo la cognición transmodular. Los partidarios de la posición eliminatoria saben que hay otras clases de estados cognitivos y de procesos tales como los que implican el pensamiento, la resolución de problemas, la deliberación o los actos de lenguaje, estados que no son transmodulares. Son centrales. El eliminador puede entonces elegir entre minimizar la significación de la cognición central o bien sostener que la cognición central podría caracterizarse y explicarse (reducirse) de arriba abajo (*bottom up*) en los términos de la cognición transmodular, aumentada con algunas hipótesis. La mayor parte de los partidarios serios de la eliminación, cualquiera que sea su concepción del espíritu, elegirán sin duda la segunda opción. Si ese proyecto de explicación de arriba abajo se concretara, el análisis cognitivo teórico podría extenderse a la cognición central.<sup>6</sup>

Aquí lo que debe servir de test es el concepto de información.

La cognición es una empresa que se refiere al tratamiento de la información. El concepto de información está en el centro del conflicto entre la posición eliminatoria y la psicología del sentido común. El proyecto eliminatorio sólo puede alcanzar éxito si la información que opera en la cognición central puede caracterizarse y explicarse en términos transmodulares. Las dos secciones que siguen tratan de dar una respuesta a esta cuestión.

## Un nuevo examen del concepto de información

### 1. "Actitudes proposicionales"

El argumento en favor de la posición eliminatoria y contra la psicología del sentido común se funda en la noción capital de actitudes proposicionales. El concepto de proposición es semántico, pues se supone que representa los aspectos semánticos o referenciales (*aboutness*) de la información manipulada en la cognición. Las propiedades semánticas de la información son reales e importantes pero, ¿por qué debemos suponer que ellas constituyen la forma última de la información que guía nuestra cognición y nuestra conducta? La información presente en la cognición no es ni exclusivamente semántica ni plenamente operante en virtud tan sólo del hecho de ser semántica. Esa información tiene también propiedades físicas y propiedades sintácticas, así como otras propiedades. Cada tipo de propiedad está ejemplificado en un nivel cualquiera. Y sin embargo ni el nivel físico ni el nivel semántico son necesariamente los niveles (últimos) en que se forman las actitudes y arrancan las conductas. Una computación formal dentro del espíritu es física, pero sabemos que un análisis físico no nos dirá lo que es esa computación ni lo que ella hace en el espíritu. Lo mismo cabe decir de una codificación semántica de la información y de su estructura semántica: un análisis de la segunda puede no revelar ni la naturaleza ni la función cognitiva de la primera. Aquí hay un paralogismo muy difundido en la filosofía contemporánea del espíritu (paralogismo que no se basa ni en un argumento ni en un valor explicativo), que consiste en sostener que la forma semántica de la información (es decir, la información codificada en condiciones semánticas e intencionales) entra en posiciones funcionales en su condición de *semántica* y engendra conductas. Esto es totalmente falso.

Antes de ver por qué es falso, detengámonos a considerar cómo el paralogismo semántico presta ayuda a la posición eliminatoria. Si la información cognitiva fuera totalmente analizable como información semántica, entonces su hábitat sería la cognición transmodular (porque es allí donde se constituye la información semántica) y la información cognitiva estaría sometida a condiciones transmodulares. Los procesos implicados serían los que las ciencias de la cognición pueden comprender y explicar por entero. La situación es por consi-

guiente clara: si se acepta el paralogismo semántico, incorporado en el concepto de actitud proposicional, se acepta ya una buena parte de la argumentación eliminativa.

¿Qué es exactamente una actitud proposicional? Este concepto, atribuido a la psicología del sentido común, es en realidad una elaboración filosófica. El sentido común se apoya en los conceptos más específicos de creencia, de deseo, de intención, etc. Los filósofos abstraen de estas nociones específicas una forma común que, desde Russell, se concibe como una actitud en relación con un contenido semántico. Aunque la actitud proposicional se analice de manera diferente, digamos, como una frase o una fórmula sintáctica o alguna otra forma de representación, la intención analítica es siempre la de representar los aspectos *semánticos* de la información codificada en una representación. Es esa intención semántica del análisis filosófico lo que debe ponerse en tela de juicio. Las preguntas que debemos hacernos son: ¿es la noción semántica de actitud proposicional una explicación conveniente de los conceptos del sentido común? ¿Tiene la noción semántica de actitud proposicional un sentido explicativo independiente? Para responder a estas preguntas sin caer en una petición de principio, debemos establecer algunas distinciones.

Para comenzar, adoptemos una actitud neutral y llamemos *actitud informacional* al concepto que estamos investigando. En una actitud informacional hay dos componentes principales, la actitud y la información. Cada uno de ellos puede caracterizarse en diferentes niveles de abstracción. Como la información es el concepto crítico de nuestra discusión, dejaré entre paréntesis las cuestiones referentes a las actitudes. Hago la hipótesis de que todos nosotros somos seres hechos para tener las actitudes que tenemos y que éstas entran en interacción de maneras igualmente predeterminadas. Aunque dispusiéramos de un gráfico científico completo de la “mecánica” de la formación y de las interacciones de las actitudes, sostengo que los objetos de esas actitudes (la información) son lo más importante para la psicología del sentido común.<sup>7</sup>

Retornemos pues al concepto de información. En la cognición hay dos aspectos de información, uno interno, el otro externo. La distinción intuitiva es la siguiente. La información es una estructura o una codificación producida en un organismo por una entrada que tiene su origen en una fuente exterior. Un

aspecto de la información es *interno* (y por consiguiente cognitivo) sólo si está codificado de cualquier forma en un estado de la composición material del organismo, y es esa codificación lo causalmente eficaz en la cognición. Un aspecto de la información es *externo* si produce cognición (como un estímulo luminoso) pero no es interno a ella. Llamo *semánticos* a los aspectos o condiciones de información externos tales como la referencia y las condiciones de verdad que caracterizan el hecho y el grado de interacción que hay entre las codificaciones internas de la información (digamos representaciones) y sus objetos externos. Considero que esas interacciones y diversas reglas sociales, convenciones, prácticas y factores contextuales son condiciones de la información semántica. Por fin, trato los conceptos, las significaciones, los prototipos, las estructuras de reconocimiento, etc. como aspectos *intencionales* de la información, aspectos internos de la cognición. Los aspectos intencionales están o bien explícitamente representados en la memoria como estructuras de datos, o bien son implícitamente utilizados como reglas y procedimientos adquiridos o innatos. Lo importante es que, aunque dichos aspectos obran como condiciones sobre la información (en el sentido en que ellos representan hechos y representaciones semánticas), las dimensiones intencionales operan sobre la información de manera cognitiva, interna. En consecuencia, una información semántica está *a la vez* regida por condiciones internas y externas. Es esta doble contribución lo que determina aquello a lo que se refiere la información particular.

El concepto semántico habitual de actitud proposicional implica que, cuando una información particular llega a ser el objeto de una actitud, son sus propiedades *semánticas* las que constituyen el objeto de la actitud y las que llegan a ser funcionalmente eficaces. Cuando (motivado por lo que informa el *New Yorker*) llego a creer que, habiendo sido invitado a ocupar una posición académica en Princeton, George Washington se preguntaba cuánto tiempo tardaría en obtener su nombramiento de titular, la concepción usual de las actitudes proposicionales implica que mi creencia refleja y pone en juego solamente los aspectos semánticos de información desarrollados en mi cabeza en esa ocasión. En otros términos, la tesis usual dice que la información que es el objeto de mi creencia puede explicarse por entero como un agregado de parámetros internos y externos que

fijan su carácter referencial semántico. La tesis usual nos dice que las relaciones y las leyes principales de las actitudes proposicionales son de naturaleza semántica porque son sensibles a la preservación de la verdad (lo cual es una virtud semántica). Se supone que las relaciones capitales de contenido entre las creencias son lógicas y por consiguiente semánticamente sensibles en la forma de la implicación, en la forma de la equivalencia, etc. Las leyes que gobiernan las relaciones entre creencias tienen una forma normativa y son de nuevo semánticamente sensibles (consistencia, coherencia, término de una relación deductiva). Y así sucesivamente. Esta es la teoría estándar, tan dominante y tan errónea.

## 2. Interpretaciones de la información

Los partidarios del programa eliminatorio explotan hábilmente la interpretación semántica de las actitudes informacionales al atribuirlo todo primero al sentido común, luego al mostrar que esa actitud es, por un lado, explicativamente inútil si se la concibe de manera puramente externa y, por otro lado, está sujeta a la explicación cognitiva teórica, si se la concibe de manera puramente interna. Es fácil jugar la primera mano de la partida. La interpretación externa de la información semántica no puede por sí misma explicar nada de lo cognitivo o de la conducta. Debemos pues preguntarnos por qué el sentido común, después de tantos años de experiencias cotidianas, de retroacciones y de series de atribuciones de actitudes y racionalizaciones logradas, deba hacer la hipótesis inversa. El requisito del internalismo, que parece seguir la psicología del sentido común, queda así violado. Por eso ningún aspecto externo de la información tiene en sí mismo eficacia causal interna; y cuando un aspecto está codificado de manera interna, ya no es más semántico. Las nociones semánticas como la de referencia, como las de condiciones de verdad, etc. son utilizadas para *evaluar* las maneras en que las codificaciones internas de la información se comportan cuando se las compara con sus blancos externos. Como el sentido común está en realidad constantemente empeñado en evaluaciones semánticas, es más sensato suponer que su empleo de las nociones semánticas es *evaluativo* antes que descriptivo y explicativo en el caso de la cognición. Volveré a tratar este punto.

La segunda mano de la partida interpreta las actitudes informacionales de manera internalista. Hay por lo menos tres lecturas internas de la información: según la composición material de los organismos, según condiciones sintácticas y según condiciones intencionales. La interpretación por la composición material es demasiado débil para dar una individuación de la información cognitiva, pues no logra especificar su codificación formal ni su papel funcional. La codificación así como el papel son, en principio, compatibles con numerosas clases de composiciones materiales. Este argumento es muy corriente, de manera que no necesitamos detenernos a considerarlo aquí. Podemos observar también que el sentido común no se ocupa tampoco gran cosa de la composición material. Dejemos pues estas cuestiones de lado. La interpretación sintáctica da los mecanismos de la representación interna de la información. Muchos, e incluso eliminadores (por ejemplo Stich 13), piensan que la lectura sintáctica es la interpretación *última* de la información. Trátase del paralogismo sintáctico, el cual implica que no hay otro nivel de abstracción en el que podamos de manera fructífera (es decir, con un fin explicativo) individualizar las propiedades internas de la información. El paralogismo sintáctico sirve a la causa eliminatoria mejor aún que el paralogismo semántico porque la sintaxis de la representación cognitiva, a diferencia de su semántica, está enteramente de acuerdo con su semántica y con el solipsismo metodológico, por lo cual es en consecuencia un objeto legítimo de la explicación cognitiva teórica. El sentido común nada esclarecedor tiene que decir sobre este punto y, por supuesto, no le preocupa. Si la información cognitiva es solamente un asunto de composición material y de sintaxis, el sentido común no tiene dominio alguno conceptual sobre la información, de manera que se lo puede abandonar.

Y sin embargo, la interpretación sintáctica tampoco puede fijar la información cognitiva. Una razón superficial y familiar es la de que la misma estructura sintáctica (*Por todas partes hay peras en el terreno universitario*) puede codificar contenidos semánticos diferentes, así como diversas estructuras sintácticas (por ejemplo en diferentes lenguas) pueden codificar el mismo contenido semántico. Hasta una sintaxis sublingüística más fina tropezará con el mismo problema, como habremos de verlo enseguida. La razón más seria es la de que la sintaxis cognitiva

debe codificar información en formas que, sin dejar de respetar los aspectos semánticos de los intereses del organismo, guíen la conducta. Este trabajo semántico y funcional de la sintaxis no puede explicarse desde el punto de vista sintáctico sin recurrir a los factores que lo motivan.

Para ver aquello de que además tenemos necesidad, formulemos una pregunta artificiosa: ¿cuáles serían las condiciones en las que el paralogsimo sintáctico podría convertirse en una verdadera concepción del espíritu? Serían las condiciones en las que la interpretación sintáctica contara como la interpretación última de la información, interpretación que fijaría la forma de la información al regir la cognición y la conducta de un organismo. ¿Qué clase de organismo podría hacer correcto el paralogsimo sintáctico? Pues bien, allí están los ordenadores para ayudarnos. Consideremos uno de ellos al que llamaremos HAM y que nunca sabe lo que hace falta. HAM es un ordenador cuyos mecanismos sintácticos registran un día los vuelos de los aviones y al día siguiente verifica los nombres del personal de la compañía y si están suscritos a publicaciones libertarias. Somos nosotros quienes aseguramos la interpretación semántica de lo que hace HAM tanto hoy como mañana. Todo cuanto hace HAM es calcular. Calcula pero no representa nada. No puede *representar* ninguna clase de cosas o en todo caso representa lo que decimos nosotros.

¿Qué implica la diferencia que hay entre computación y representación? ¿Qué es lo que hace que la información manipulada por HAM sea sintáctica pero no semántica? Podemos formularlo del modo siguiente. La razón por la cual HAM puede trabajar sobre los planes de vuelos es la de que ha sido programado de manera tal que sus representaciones sintácticas y sus operaciones codifiquen y preserven los aspectos semánticos pertinentes en el dominio de interpretación apuntado (los planes de vuelo). Esta correlación o coincidencia sintáctica y semántica no es algo que HAM conozca. Nada se le ha dicho a HAM sobre esta correlación, y él mismo no está dotado de los recursos internos que le permitan registrar o captar la correlación alineando sus formas y operaciones sintácticas sobre los aspectos semánticos del caso. Es la realización interna (en una forma cualquiera) de la alineación de la sintaxis sobre la semántica lo que representa toda la diferencia en la cognición, realización de

la que HAM está privado. Esto hace de HAM un motor puramente sintáctico.

Podemos enunciarlo en otros términos. HAM tiene la aptitud sintáctica de calcular planes de vuelos pero carece de *otra* aptitud que consiste en individualizar los planes de vuelo *como tales* y en distinguirlos, por ejemplo, de los legajos del personal o de cualquier otra cosa. El pobre HAM no tiene *conceptos*. Por eso, las formas sintácticas calculadas por HAM no representan nada en particular y por consiguiente no tienen ninguna *significación*. Se puede decir que una forma sintáctica tiene una significación sólo si ella se desarrolla en la forma de ciertos conceptos o de otras rutinas de clasificación destinadas (por el aprendizaje, la evolución, la intervención divina o lo que fuera) a proyectar formas sintácticas sobre aspectos semánticos. Leer una sintaxis que expresa un contenido a la luz de conceptos y de significaciones es leerla *intencionalmente*. En este contexto los conceptos pueden considerarse como funciones que envían formas sintácticas a aspectos semánticos, como condiciones que aseguran, en primer lugar, que la codificación sintáctica de la información sea sensible a los aspectos semánticos de la significación para el sujeto de la cognición y, en segundo lugar, que la computación de la información preserve esos aspectos semánticos. Los conceptos tienen esta función porque deben asegurar que la conducta del organismo esté dirigida a diversos blancos del ambiente. Esto es terriblemente enigmático pero supongo que bastará, especialmente si nos permite responder a nuestra segunda pregunta.

¿Qué significa todo esto en relación con la posición eliminatoria, con el sentido común y con las actitudes informacionales? Los aspectos intencionales de la cognición, los conceptos y las significaciones en particular no son ni tienen por qué ser exclusivamente asunto de la psicología del sentido común. El sentido común ni siquiera se preocupa por la naturaleza y función de los conceptos y significaciones de base que empleamos corrientemente, a menos que en ellos falle algo. Esto sugiere que la psicología del sentido común no es realmente una teoría de nuestro programa intencional básico. Pero el programa intencional *puede* explicarse de manera teórica cognitiva. Las teorías psicológicas del aprendizaje estudian la formación de los conceptos y sus aplicaciones. Las teorías de la visión y del



lenguaje estudian cómo la sintaxis hace entrar en acción los aspectos semánticos cruciales de la visión y del lenguaje, etc.

Allí donde la ciencia tiene un asidero, el sentido común suelta la presa. Si la psicología del sentido común no tiene en principio interés por nuestra ejecución intencional y si ésta puede explicarse por la ciencia de la cognición, deberíamos inferir sensatamente de ello que el sentido común no puede formar parte de lo que procede del dominio de lo intencional sin correr el riesgo de ser eliminado. Y si el sentido común insiste en desempeñar un papel psicológico propio, deberíamos inferir sensatamente que ese papel se sitúa en *otra parte*. Estas no son inferencias fáciles de hacer, desde luego, ni siquiera para los filósofos que defienden el sentido común, pues también ellos piensan que si el sentido común explica y esclarece algo del espíritu, ese algo es precisamente su intencionalidad. (El paralogsimo semántico es el pan cotidiano de la filosofía analítica contemporánea.) Y yo rechazo esta idea pues sostengo que la intencionalidad (o la aptitud referencial mental) es apenas el comienzo de la imagen de lo mental cuyo vehículo son los sentidos.

Si la intencionalidad no es principalmente asunto del sentido común, ¿qué es asunto de éste si debe ocuparse de la información cognitiva en aspectos internos? Si la información registrada por un estado mental no puede fijarse de manera separada o conjunta, ya por sus condiciones semánticas externas, ya por sus parámetros de composición material, de sintaxis o de intencionalidad, entonces, ¿qué es lo que en el espíritu se ocupa de estas cosas?

### 3. La información mental

Evidentemente en la información cognitiva hay más de lo que dice la sabiduría filosófica del momento. O por lo menos esto es lo que quiero sugerir. En otros lugares (1, 2 y 3) digo algo más sobre este particular. Aquí sólo se encontrará un corto resumen. Ya introduje antes una nueva forma de información llamada información *mental*. Esa forma tiene una codificación interna, pero es de naturaleza pragmática. Es esa forma de codificación, y no sus rasgos sintácticos y semánticos, la que es funcionalmente eficaz en la cognición y en la conducta. Por eso los paralogsimos semánticos y sintácticos son falsas representaciones de la verdadera cognición.

Para comprender lo que esto quiere decir, comencemos por distinguir la noción de información mental de otras nociones familiares e importantes de información. En primer lugar debemos distinguirla del concepto de *información de entrada* (*input information*) tan bien analizado recientemente por Fred Dretske, (9) Este último concepto se refiere a las interacciones entre las fuentes externas y los organismos, específicamente a la información entregada por señales físicas en umbrales sensoriales. Se trata de una forma de información *física* codificada en estructuras materiales. La información de entrada no especifica la manera en que los organismos se representan las entradas ni cómo esa representación produce otras representaciones y acciones. Para esto tenemos necesidad de la forma *semántica* de la información. Sensible a condiciones sintácticas e intencionales, esta forma se define en aspectos semánticos que interesan al organismo y especifica cómo esos aspectos se adaptan a las condiciones del caso. En particular, la información semántica codificada por una representación puede especificar qué estados de cosas son compatibles con esa representación y cuáles no lo son. Esto aproxima mucho el concepto de información semántica al concepto de significación y lo hace caracterizable hasta el nivel intencional de abstracción.

Hemos visto que esto bastaría para comprender la cognición. Pero también debemos determinar cómo las posibilidades semánticas disponibles en virtud de la información de entrada e identificadas por sus codificaciones representacionales afectan, por ejemplo, la incertidumbre en que se encuentra un agente cognitivo frente a una situación, cómo afecta lo que él ya sabe de esa situación, cómo la incertidumbre se reduce y cómo el agente obra partiendo de una nueva información. Si cambiamos esta historia de incertidumbre por el acto de resolver un problema, de tomar una decisión, de escoger una hipótesis, de comunicar algo o de realizar una acción, el análisis de base es idéntico. En todos los casos se trata de un análisis en el que la información semántica disponible por obra de la entrada y codificada en condiciones sintácticas e intencionales queda *relativizada*, en nuevas condiciones, al estado cognitivo del momento del agente cognitivo. El resultado es información mental.

Consideremos por ejemplo el caso de la comunicación. Oigo una enunciación. Su significación me permite representarme sus posibilidades semánticas o sus condiciones de verdad. Pero

yo no obro sobre la base de únicamente significaciones. Espero a que una enunciación haga *otra* contribución a mi cognición al entrar en interacción con lo que sé, con lo que ignoro, con lo que espero y proyecto sobre una situación. Es en esta interacción *suplementaria* (que va más allá de la semántica y de la intencionalidad) cuando la información mental cobra forma; y cobra forma pragmáticamente y de manera aumentativa, según el estado cognitivo agregado en el que me encuentro y según los objetivos de conducta que persigo.

Los hechos que se refieren a la información mental son problemáticamente pero significativamente familiares. El lector y yo podemos encontrarnos en una situación en la que, al recibir (tipos de) entradas similares, visuales o verbales, reaccionemos de manera diferente. Esas diferencias cognitivas o de comportamiento pueden existir aun cuando los aspectos claves de la información fijados por esos tipos de entradas similares sean también tipos similares. Imaginemos, por ejemplo, que el lector y yo oímos el malicioso rumor de que Robert Vesco propuso reconstruir el programa de baloncesto de Tulane. Podemos hacer la hipótesis plausible de que fonética, gramatical y conceptualmente los dos tratamos este enunciado de entrada de maneras similares. Suponiendo que nuestros conceptos de Vesco, de reconstrucción, de baloncesto y de programa sean semejantes, ambos leemos la misma significación literal en el enunciado. Y sin embargo, después de haber calculado esa significación, el lector dice: "¡Ah! ¡Ah! ¡Qué interesante!", en tanto que yo veo luces rojas y dientes rotos. No son pues exactamente las mismas inferencias, evaluaciones y conductas.

¿Por qué esa diferencia? Lo que ocurre es que yo sé más sobre el programa de baloncesto. También puedo tener expectativas y aprensiones diferentes de las del lector sobre esta cuestión del baloncesto. La diferencia entre nuestras expectativas y aprensiones puede llegar a explicar esa diferencia de conductas. Pero *antes* debe haber una diferencia de información que explique toda una cantidad de acciones cognitivas (inferencias efectuadas, creencias fijadas) que yo cumplo y que el lector no cumple. Si esto es así, la información que el lector y yo extraemos de la información de entrada será diferente. Se trata de información mental.

Imaginemos que nuestro HAM esté dotado de analizadores perceptuales, de reglas formales y de conceptos que fijan las

significaciones. HAM puede tener acceso a una información de entrada  $p$  referente a una situación  $s$ , puede representar  $p$  sintácticamente en una fórmula  $f$  y puede hacer que  $f$  llegue a representar conceptos apropiados que fijan el sentido de  $f$ ,  $m$ . En esta cuestión, HAM registra, calcula y representa una información de entrada hasta el nivel intencional. Hasta aquí HAM puede imitar los procesos perceptuales, sintácticos e intencionales por los que pasamos nosotros cuando comprendemos una frase. Supongamos ahora que en el caso de HAM, a diferencia del nuestro, la salida intencional  $f(m)$  no vaya a ninguna parte, pues HAM no tiene memoria ni conducta a la que pudiera servir su memoria. Desde el momento en que calcula y representa un  $f(m)$  cualquiera, ese  $f(m)$  desaparece pues no entra en interacción con ningún conocimiento anterior, con ninguna creencia, con ninguna expectativa, con ningún plan, con nada. Los estados cognitivos presentes de HAM no influyen nada en lo que ocurre en la salida intencional  $f(m)$ .

Nuestro  $f(m)$  está en cambio sujeto todavía a una interacción con nuestros estados cognitivos presentes (creencias, expectativas, etc.). Como la información es una cuestión de estructuras creadas en interacción con otras estructuras, este *otro* tipo de interacción, gobernado por las nuevas formas de las condiciones (orientadas hacia la acción y pragmáticas) debería señalar la existencia de una nueva forma de información. Como no dispone de los recursos necesarios para realizar semejante interacción interna, HAM es incapaz de fijar la información mental que  $f(m)$  habría agregado a la configuración presente de las creencias almacenadas por HAM. Pero en HAM no hay ninguna configuración de este tipo, de manera que no se agrega ninguna información ni se fija ninguna creencia. Ocasionalmente cada uno de nosotros puede asemejarse a HAM cuando, un poco desfasados, registramos una entrada aislada, por ejemplo, una enunciación que podemos comprender en el nivel intencional (o de la significación), es decir, como  $f(m)$ , pero no más de eso porque no podemos relacionar nuestros conocimientos o nuestras expectativas con la significación de la enunciación. El concepto de información mental se propone precisamente captar la diferencia que hay entre lo que está representado, como significación o como  $f(m)$ , y lo que esa representación agrega a lo que ya sabemos, esperamos o anticipamos.

¿Por qué es importante todo esto? ¿Por qué es importante la información mental? La información, según vimos, se presenta siempre en la forma de una estructura engendrada por interacciones con otras estructuras obedeciendo a leyes y condiciones apropiadas. La razón de que la información semántica sea distinta de la información mental es la de que la primera está engendrada por entradas en condiciones materiales, sintácticas e intencionales, mientras que la segunda está engendrada por la información semántica con los estados cognitivos y de conducta presentes en nuevas clases de condiciones (de naturaleza pragmática e inductiva debida a lo que está almacenado, a lo que se cree, se proyecta, se espera, etc.) La información mental tiene una función pragmática que consiste en vincular la información semántica con la cognición y con la acción presentes del organismo. La información semántica, fijada a la vez por interacciones causales con el ambiente y por condiciones sintácticas e intencionales, no puede por sí sola hacer este trabajo. Para emplear una metáfora, digamos que la información semántica se vuelve hacia atrás, hacia el ambiente para representarlo, en tanto que la información mental toma la información y la proyecta hacia adelante para ver cómo puede servir a la cognición y a la acción presente y futura. La información semántica está condicionada por relaciones de mundo y espíritu; la información mental está condicionada por relaciones de espíritu y acción.

La fijación de la información mental es un proceso inductivo en el que diversos estados que codifican datos entran en interacción con una representación -blanco apuntado (la salida intencional presente) y fijan una nueva información en el conocimiento del agente cognitivo atendiendo a lo que éste ya sabe, espera y quiere hacer. El proceso de fijación debe ser holístico, *ad hoc* y contextual y por consiguiente en alto grado individualizado a causa de estas interacciones. Síguese de ello que las condiciones de estos procesos mentales inductivos no serán invariantes, a diferencia de las condiciones sintácticas e intencionales. Los procesos que fijan la información no pueden ser invariantes porque están subdeterminados no sólo por el contexto sino también por las coacciones externas de la composición material, de la sintaxis y del nivel intencional. La fijación de la información es un proceso auténticamente central y pragmático.<sup>8</sup>

## 5. Las actitudes mentales brevemente reseñadas

A fin de sacar las conclusiones de nuestra teoría de la información mental a los efectos de comprender la psicología del sentido común, debemos ante todo poner en relación esta teoría con nuestra concepción de las actitudes mentales. El paralogsismo semántico quiere que las actitudes mentales tengan contenidos semánticos como objetos. Esto no significa gran cosa. Las actitudes mentales deben tener relación con la manera en que la información anima y mueve la cognición y la conducta presentes de cada uno. Y, como ya vimos, no es la información semántica como tal la que anima y mueve. Creemos y deseamos porque debemos obrar y para obrar debemos registrar la información que relaciona la acción con nuestro estado cognitivo presente. En otras palabras, tratamos la información atendiendo a la creencia y al deseo porque esa información debe ponerse al servicio de nuestra acción y de nuestra cognición del momento.

Este es un hecho teleológico craso. HAM no tiene deseos ni creencias pero no tiene la obligación de obrar, porque puede tratar la información en el nivel semántico. Verdad es que, por muy buenas razones, *tratamos* a menudo las creencias y otras actitudes como relaciones (corriente arriba) con contenidos semánticos solamente e ignoramos sus relaciones (corriente abajo) con la cognición y con la acción presente y futura. Pero ésta es sólo una táctica oportunista que no nos dice lo que son las creencias y las otras actitudes. Puedo tratar un libro como si fuera un martillo, pero eso no hace del libro un martillo. Exploto las propiedades que el libro tiene para hacer entrar clavos en la pared, pero ignoro las demás propiedades que esencialmente tienen los libros. Las dimensiones semánticas de una creencia, la información semántica o el contenido que ella codifica pueden ser fundamentales en un contexto, por ejemplo, el de la evaluación de la verdad o el de la interpretación intencional. En esos casos una creencia puede ser *tratada* como una relación con un contenido semántico (digamos, una proposición). Por razones biosociales importantes, el sentido común es sensible a la interpretación semántica e intencional y con frecuencia trata las creencias de conformidad con esta circunstancia. Pero no está dentro de la naturaleza de una creencia el ser sólo semántica, pues de otra manera su función cognitiva queda oscurecida.

Pudiera parecer que el problema se reduce tan sólo a una cuestión de terminología. Podemos llamar “creencia” la simple actitud intencional frente a un contenido semántico y llamar cualquier otra cosa a la actitud frente a la información mental. Sin embargo, esta maniobra terminológica no debe suponer que el espíritu cognitivo registra y opera con las *dos* clases de actitudes, pues de otra manera nos encontraríamos con dos problemas de coordinación inútiles. Uno es el de la coordinación entre la mitad semántica/intencional y la mitad mental del concepto de “creencia”; el otro es el de la coordinación entre todo lo que cuenta como creencia semántica/intencional y las demás especies de actitudes, tales como deseos e intenciones. (Este último punto presenta un problema, porque los deseos y las intenciones se refieren a ella. Cuando una creencia intencional está orientada hacia la acción, dicha creencia ya no es sólo intencional, es mental.)

En la vida real de la cognición los dos problemas de coordinación tienen una solución simple y natural sólo porque la inscripción de las actitudes mentales es real y eficaz. La actitud intencional frente a un contenido semántico no tiene por sí misma realidad cognitiva ni eficacia; es tan sólo el artefacto prácticamente motivado de la atribución (luego vuelvo a considerar este punto) o el artefacto filosóficamente motivado del análisis fenomenológico y metafilosófico. En esta instancia cognitiva natural, un sujeto cognitivo difícilmente pueda tener una actitud frente a un contenido semántico de manera solamente intencional, sin hacerle desempeñar al mismo tiempo otro papel cognitivo y de conducta. Una vez que el contenido desempeña ese papel ya no es solamente semántico y la actitud ya no es solamente intencional pues su estructura misma y su operación están reguladas ahora por otras condiciones, no semánticas y no intencionales.

Para resumir, mi estrategia ha sido primero dejar que los hechos y regularidades mentales condicionen la información, así como dejamos que los hechos y regularidades intencionales condicionen las codificaciones y operaciones sintácticas subyacentes; y en segundo lugar, declarar cognitivamente real y eficaz el tipo de informaciones y actitudes que están inscriptas y condicionadas en última instancia. Para poder llegar a tratar tan correctamente la pragmática de la cognición y de la conducta, el sentido común debe reconocer y explotar esta verdad

referente a la información mental y a las actitudes. Esto es lo que me falta mostrar como conclusión.

## 6. La psicopráctica del sentido común

Si el sentido común interpreta la cognición y la conducta de un agente atendiendo a la información mental en la forma de actitudes, las estrategias que emplea para ese fin deben ser individualistas y sin embargo realistas, porque esto forma parte de la naturaleza misma de la información mental: ésta es cognitivamente real y causalmente eficaz por más que esté constituida en condiciones esencialmente individualizadas y pragmáticas. Creo que no se necesita ninguna argumentación complicada para mostrar, primero, que el sentido común se interesa (entre otras cosas) por la explicación de la cognición y de la conducta individuales desde el punto de vista de la información mental en forma de actitudes y, segundo, que el sentido común es capaz de presentar historias individuales sobre la manera en que los agentes particulares fijan la información, a la cual extienden sus actitudes cognitivas y no cognitivas y partiendo de la cual obran. La fijación de la información mental es esencialmente (en oposición a trivialmente) individual y pragmática, porque se refiere al estado cognitivo y de conducta del agente en la forma de incrementos efímeros. La forma de la información es aumentativa porque, según vimos, las entradas encuentran siempre conocimientos almacenados en la memoria, lo cual hace que el incremento nuevo agregado al antiguo sea cognitivamente pertinente y eficaz. Sin embargo ese aumento es reconocido por el sujeto cognitivo sólo dentro de un contexto temático particular, dentro de cierta forma conceptual en relación con otros incrementos compatibles con lo que el sujeto sabía, con aquello de que estaba inseguro o con aquello que esperaba. El incremento es efímero porque todos sus parámetros (los que acabamos de enumerar) cambian de valor de un contexto a otro.

Reunamos ahora nuestros resultados. Tenemos hechos verdaderos e importantes sobre la cognición en la forma de la información mental de las actitudes. Esta información dirige la cognición y la conducta de manera esencialmente pragmática; y esto es lo que hay que comprender si se desea comprender la cognición. Tenemos también un paradigma sorprendentemente

vigoroso de conceptualización y de explicación de la conducta, paradigma que llamamos la psicología del sentido común, cuyos intereses explicativos parecen dirigirse a la información por las actitudes y cuyas estrategias explicativas parecen adaptarse a su objeto. Podemos recordar aquí que existe una hipótesis metodológica plausible, tratada en el párrafo 3, según la cual la naturaleza ontológica del objeto de una teoría dicta en gran medida la manera en que el objeto es conceptualizado y explicado. Teniendo en cuenta todo esto es natural formar la hipótesis de que las estrategias individualistas del sentido común están reguladas a fin de adaptarse a la naturaleza de un objeto de explicación esencialmente individual y pragmático. Si esta hipótesis es plausible, entonces la posición eliminatoria debe estar equivocada por lo menos sobre una forma de lo mental y sobre su localización, la psicología del sentido común.

Y sin embargo, por plausible que parezca el argumento, no será muy convincente si no tratamos de analizar un poco más la psicología del sentido común misma. Esta será nuestra última tarea. Recordemos antes una observación que hicimos en el párrafo 3, según la cual nos son familiares los hechos referentes a la información. La razón de ello está en la naturaleza misma de la psicología del sentido común. Si nos parece fácil seguir y especificar la información mental contenida en actitudes, ello se debe a que esa información es para nosotros como una segunda naturaleza. Lo que el sentido común nos ayuda a hacer tan bien, sin esfuerzo ni reflexión, es una cuestión de rutina y de práctica antes que de teoría. Es una práctica que reconoce información mental en sus diversas manifestaciones en la forma de actitudes, sin reconocer u ocuparse necesariamente de sus formas sintácticas, semánticas o intencionales subyacentes. Esto no puede ser diferente de nuestro conocimiento práctico, por ejemplo, de mesas, de sillas o de cómo hacer una tortilla, conocimiento que no se ocupa de la constitución intrínseca de los objetos.

Si comprendemos que la psicología del sentido común es una práctica implícita, irreflexiva, dirigida hacia una familia de fenómenos mentales, en lugar de ser una teoría explícita de esos fenómenos, también podemos ver por qué “la psicología del sentido común” es una expresión tan engañosa. Aquí no hay ningún *logos* pues se trata solamente de *praxis*. El sentido común es una psicopráctica, no algo en que intervenga el *logos*; es

una práctica que para funcionar se apoya en conceptos y relaciones conceptuales que representan propiedades objetivas y regularidades interesantes. Para conducir un automóvil, tengo necesidad de cierto número de conceptos (velocidad, aceleración, freno, etc.) y de cierto número de correlaciones interconceptuales sobre variadas regularidades, por ejemplo, hacer funcionar el arranque en posición de estacionamiento pone en marcha el motor, frenar violentamente en superficies mojadas hace patinar el automóvil, etc., todo en un nivel superficial de descripción. Mi conocimiento de los automóviles y su manejo es algo práctico, no teórico. No estudio los automóviles, me contento con hacer algo con ellos fundándome en un conocimiento práctico mínimo. Ese conocimiento tiene una base realista pues se refiere a los automóviles y a las situaciones en que éstos operan. De manera aun más implícita, esto es cierto en el caso de andar en bicicleta, de nadar o aun de caminar. Asimismo los gatos no podrían atrapar ratones si sus prácticas de caza no estuvieran guiadas por nociones realistas sobre el comportamiento de los ratones. Esas nociones no necesitan formar una teoría del comportamiento de los ratones. Lo mismo ocurre con el sentido común: para hacer correctamente su trabajo debe captar aspectos explicativamente reales y causalmente eficaces de la cognición central. Si una práctica no fuera realista, ¿para qué serviría?

Y sin embargo captar la fijación de la información, representarse las aptitudes frente a la información y por consiguiente sus productos cognitivos y de conducta sólo parece ser una de las funciones del sentido común. Otra función (cuando es necesario) es la interpretación intencional, es decir, el intento de representarse el repertorio conceptual de otro sujeto cognitivo del mismo grupo o de otro grupo. La traducción y la interpretación radical, tales como las han entendido los filósofos desde Quine en adelante, son ejemplos dramáticos de interpretación intencional. La tarea es aquí igualmente psicológica, pero incumbe a una forma de competencia cognitiva: nuestro programa conceptual. En cambio, la psicopráctica, que se ocupa de la información mental *via* las actitudes, tiene que ver esencialmente con la manera en que se desarrolla el programa conceptual en los casos reales de cognición, en las realizaciones cognitivas. Lo que también entra dentro de la competencia del sentido común es, como hemos visto, la evaluación semántica de

los estados mentales en relación con sus ambientes naturales y sociales. Sin embargo esto ya no es un trabajo psicológico, aquí hay evaluación epistémica. Todas estas prácticas distintas del sentido común se concentran en los variados aspectos de nuestra vida cognitiva. Las prácticas reflejan una sensible división del trabajo motivada por factores biosociológicos. Aquí no puedo desarrollar este punto. La manera en que llegamos a aprender, a dominar y a ejercitar esas prácticas continúa siendo un misterio. De la instancia psicopráctica de la fijación de información mental pasamos fácilmente a la instancia semántica de identificar condiciones de verdad de esa información o a la instancia intencional de la aserción de los conceptos utilizados. Y así sucesivamente.

Los filósofos que consideran que el sentido común y el lenguaje ordinario definen implícitamente diversos aspectos de la vida mental caen a menudo en el error de tomar una práctica particular del sentido común como si ella representara el sentido común en general. Consideran esa práctica específica como algo que incorpora la comprensión general de lo mental por el sentido común. Por numerosas razones, la evaluación semántica y la interpretación intencional son los casos favoritos de esos filósofos. En filosofía del espíritu y del lenguaje definen, por ejemplo, lo que son las actitudes, y por eso han llegado al concepto semántico o intencional de actitud proposicional, lo cual termina por hacer pensar equivocadamente como si se explicara así toda la parte cognitiva del espíritu.

La argumentación de este artículo en favor de las actitudes mentales y del sentido común y en contra de la posición eliminatoria se ha concentrado en un tipo de práctica del sentido común, una práctica psicológica relacionada con la información mental y las funciones de las actitudes. Es esa práctica lo que, contra la posición eliminatoria, justifica una interpretación realista pero también individualista y pragmática de las actitudes mentales frente a la información. Según esta interpretación, parece que la fijación de las actitudes mentales y su eficacia causal son hechos reales que deben tratarse atendiendo a las actuales formas de las explicaciones teóricas cognitivas. Y sin embargo no se sigue de ello que haya una competencia entre el sentido común y las ciencias de la cognición cuando se trata de comprender, en definitiva, las cuestiones humanas.

O bien las ciencias nos dan esta comprensión o bien no nos la dan porque no pueden hacerlo. En el segundo caso, no podemos proponer el sentido común para desempeñar ese papel. Lo que hace el sentido común es identificar los fenómenos que debe representarse por razones que le son propias, aun cuando esos fenómenos permanezcan sin explicación científicamente. Sólo cuando una posición filosófica como la postura eliminatoria nos dice que la psicología del sentido común es falsa porque no hay actitudes, tenemos un conflicto entre la ciencia y el sentido común. Pero el conflicto no se refiere a la explicación; se refiere a la ontología de la cognición.

La psicología del sentido común es una práctica muy vigorosa que no tiene interés programático aparente por las funciones internas de la cognición. Si mi argumentación general es correcta, esta superficialidad no deja de tener sus razones. El conocimiento de los mecanismos internos (gobernados por leyes) de la cognición hasta un nivel intencional no nos ayudará a identificar ni a explicar las actitudes mentales. Hemos mostrado que esto en principio es cierto, *independientemente* de lo que pueda decirnos el sentido común. Y sin embargo, parece que el sentido común reconoce implícitamente que la cognición central tiene facultades para fijar la información, facultades que, por un lado, están subdeterminadas por los mecanismos internos y por las leyes de la composición material en su realización formal e intencional (de ahí su superficialidad) pero que, por otro lado, son enteramente sensibles a una variedad de datos que no están regidos por leyes, datos pragmáticos que tenemos en la cabeza y fuera de ella (de ahí el aspecto práctico, múltiple, contextual, improvisado del sentido común). Si las realizaciones debidas a esas facultades, en la forma de actitudes mentales y de acciones, son lo que define y explica implícitamente una psicología individualista pero realista, entonces lo mental forma parte de lo cognitivo. Lo mental especifica los estados de la cognición central en la forma de pensamientos, de intenciones y de inferencias.

Lo mental se manifiesta, pues, allí donde las ciencias de la cognición parecen incapaces de alcanzarlo. No porque el sentido común tenga razón desde el punto de vista descriptivo, sino antes bien porque objetivamente hay algo presente en la cabeza, algo que el sentido común está forzado a reconocer y a manipular en la práctica por la existencia biológica y social.

Esto es cierto, pensemos lo que queramos del sentido común. Ello no obstante, es sorprendente que de una manera implícita, irreflexiva, práctica, el sentido común haya podido saber de continuo que la composición material, la sintaxis y la semántica de la cognición no nos lo dicen todo.<sup>9</sup>

## Notas

<sup>1</sup>No nos ocupamos aquí de la naturaleza de las representaciones, análogas o digitales, simbólicas o no, por más que una decisión sobre este problema afectaría la manera en que explicaríamos en definitiva la cognición de conformidad con el nivel de abstracción que correspondiera a esas ciencias.

<sup>2</sup>(6), pág. 67. Todo cuanto digo sobre la versión realista, incluso la sinopsis de su argumento básico contenido en el párrafo siguiente, está fundado en los trabajos de Paul Churchland (5) y (6). Y sin embargo, tanto en mi sinopsis como en la discusión que le sigue hay un elemento de generalización que va más allá de la posición eliminativa neurofisiológica de Churchland. Esto se debe a que uno puede ser eliminatorio en un nivel de abstracción más elevado que el de las neurociencias, por ejemplo, en el nivel de la neuropsicología cognitiva.

<sup>3</sup>Adam Morton (12) y K. V. Wilkes (14), en particular, defienden una versión individualista. Una parte de la bibliografía wittgensteiniana y austiniana de las décadas de 1950 y 1960 contiene también prejuicios eliminatorios de tipo individualista sobre el sentido común. Las tesis eliminatorias pueden compartir varias concepciones del sentido común que aquí distinguimos. Muchas lo hacen; la de Wittgenstein, por ejemplo, puede concebirse como una mezcla de versiones individualista e interpretacional.

<sup>4</sup>La distinción entre estas formas de explicación científica se debe a Haugeland, en (11).

<sup>5</sup>Paul Churchland me ha dicho por correspondencia que consideraba la explicación científica típica y la explicación del sentido común como explicaciones referentes a hechos y a propiedades únicas e individuales. Estoy de acuerdo con él pero distingo la unicidad trivial y la unicidad esencial. Churchland toma en consideración sólo la primera. La atribuye a la psicología del sentido común porque estima que la psicología del sentido común es una protociencia.

<sup>6</sup>Paul Churchland y quizá D. Dennett parecen pensar que una teoría de la cognición transmodular y algunas consideraciones sobre la evolución y el impacto de nuestra constitución cognitiva harán el trabajo básico (7, 8). La idea principal parece ser la siguiente. La constitución de la cognición central es la responsable de las propieda-

des de la información mental y de su fijación. De manera que la explicación de la primera explicará también la segunda. ¿Cómo explicaremos entonces la construcción transmodular? La segunda, según Churchland, puede tratarse “como algo que la evolución regula (es decir, algo que, por mutación y selección fortuita, se adapta *via* las células que son excitadas y ciertos rasgos del ambiente)”. En cuanto a la construcción central, debemos suponer que la evolución selecciona, no patrones de registro y respuestas motrices apropiadas, sino antes bien estrategias de aprendizaje como tales. En otras palabras, lo que suponemos es que “el organismo se ha adaptado tan bien que parece tener una *imagen del mundo que evoluciona*, en lugar de encontrarse en la situación en la que el organismo hubiera formado chapuceramente un análogo del proceso evolutivo mismo”. En consecuencia, el aprendizaje es evolución interiorizada. Atendiendo a esta hipótesis evolucionista sobre la construcción central, Churchland llega a esta conclusión: “Aquí, en consecuencia, el trabajo de la neurobiología y de la neuroetología es hercúleo, pero se puede apostar a que la teoría de los organismos complejos se elaborará sobre la teoría más fundamental de la semántica reguladora (entiéndase modular) de los organismos más simples, siguiendo los pasos de la evolución misma (17, pág. 14)”.

En respuesta, aceptaré la hipótesis de Churchland (de que la evolución puede explicar la construcción central), pero no acepto la conclusión (de que la construcción central pueda explicar las realizaciones). Supongamos pues que Churchland tenga razón en cuanto al hecho de que la evolución regula la construcción central, según él dice. Quisiéramos hacer dos preguntas respecto de esta hipótesis. En primer lugar, *¿cómo y hasta dónde* nos conducirá la hipótesis de la regulación en los niveles de abstracción en los que debemos examinar la cognición? Y, en segundo lugar, una vez en el nivel más elevado, *¿cuánto* de la construcción central podemos explicarnos? La respuesta más optimista a la primera pregunta es: hasta el nivel intencional, allí donde están los conceptos y las significaciones. Si la cognición humana estuviera totalmente regida por propiedades y leyes intencionales (o sometida a ellas), entonces las cogniciones particulares podrían caracterizarse por entero en ese nivel y, dadas ciertas condiciones iniciales, podrían explicarse solamente por las leyes de la construcción central. Pero nuestra teoría de la información central nos mostrará que esto no es así. La competencia intencional especificada por la construcción central subdetermina la fijación y las consecuencias funcionales de la información mental. La explicación de la competencia central no es aún explicación de sus realizaciones mentales. Esto responde a la segunda pregunta. Desarrollo este punto en (4) respondiendo a Dennett.

<sup>7</sup> Hay una cantidad de leyes de actitudes que afloran en nuestras explicaciones del sentido común. Estoy pensando en leyes como *Si X teme que p, entonces X desea que no p; Si X espera que p y descubre que p, luego X está satisfecho de que p*, etc. Semejantes leyes sólo implican actitudes, otras actitudes y acción. Entre los eliminadores, Churchland (5 y 6) y otros discuten las leyes de actitudes detalladamente. Yo las llamo leyes *actitudinales* y no leyes de contenidos, porque la fuente de su legalidad se debe a las actitudes y no a los contenidos entrañados. Es una cuestión de construcción de actitudes si el temor va acompañado del deseo, cualesquiera que sean sus contenidos. Por otra parte, una ley de contenido, aun cuando ésta rijan actitudes de información, es legal a causa de los contenidos de esas actitudes. Por ejemplo, *Si S cree que p y cree también que p implica q, entonces S cree que q* es una ley informacional porque está basada en la ley lógica del *modus ponens* contenido por la actitud de creencia, cualquiera que sea la razón de ella. En una ley de actitud es lo inverso pues el contenido, dentro de la ley, según el cual *esperanza de que p + descubrimiento de que p = placer que p* mantiene una regularidad actitudinal. Hablo deliberadamente del contenido como noción semántica y no de información para hacer resaltar el prejuicio difundido en el análisis de los contenidos de actitudes.

La mayor parte de las leyes de actitudes están especificadas por nuestra construcción de base. El sentido común tiene conciencia de ello y explota esas leyes. El sentido común puede engañarse en este punto. Y sin embargo, que tenga razón o no, la psicología del sentido común no parece, *pace* Churchland, ser una teoría de las leyes de actitudes, así como no es una teoría sobre la gramática de la visión o del aprendizaje. Es más probable que a la psicología del sentido común le incumba la noción de información. Quiero pues conceder a Churchland, a Dennett y a los demás que la “mecánica” de la interacción de las actitudes y sus leyes podría llegar a ser el objeto de una ciencia de la construcción cognitiva. Concederé también que la psicología del sentido común puede asimismo equivocarse sobre esta “mecánica”, pero continúo insistiendo en el hecho de que la psicología del sentido común está ante todo interesada por la información, cuyo vehículo son las actitudes. Sin embargo, tratándose de la información, la construcción cognitiva es mucho menos útil, como lo muestra la nota anterior.

<sup>8</sup> Desarrollo este punto con más detalles en (3).

<sup>9</sup> Hace mucho tiempo escribí un antecesor de este artículo que se proponía ser una refutación de la versión realista de la posición eliminatoria, en particular de Churchland. Luego, estimulado por nuevos pensamientos y nuevas críticas, me puse a considerar la cuestión más amplia de la psicología del sentido común en oposición a la postura eliminatoria. Agradezco a Paul Churchland, a un árbitro

anónimo de *Nous* que hizo excelentes observaciones, a mis amigos de los seminarios del CNRS de París, a los que me oyeron en Helsinki y a mis amigos de Tulane y de la universidad de Nueva Orleans, Harvey Green, Carolyn Morilo y Norton Nelkin, quienes después de haber oído hablar de mis diversas radicaciones e instalaciones, pueden creer que estoy financiado por Hoi Polloi Inc. Si lo estuviera, me encontraría ahora en París.

## Referencias bibliográficas

1. Radu J. Bogdan, "Mind, Content, and Information", *Synthese*, 70 (1987), 205-227.
2. Radu J. Bogdan, "The Manufacture of Belief", *Belief*, Radu J. Bogdan (comp.), (Oxford, Oxford UP, 1986).
3. Radu J. Bogdan, *Mind and information*, de futura aparición.
4. Radu J. Bogdan, "The Intentional Stance Reexamined", *The Behavioral and Brain Sciences*, 8 (1985), 759-760.
5. Paul Churchland, *Scientific Materialism and the Plasticity of Mind* (Cambridge, Cambridge UP, 1979).
6. Paul Churchland, « Eliminitive Materialism and Propositional Attitudes », *Journal of Philosophy*, 78 (1981): 67-90.
7. Paul y Patricia Churchland, "Stalking the Wild Epistemic Engine", *Nous*, 17 (1983), 6-18.
8. Daniel Dennett, *Brainstorms* (Cambridge, MA, MIT Press/Bradford, 1978).
9. Fred Dretske, *Knowledge and the Flow of Information* (Cambridge, MA, The MIT Press/Bradford, 1981).
10. Jerry Fodor, *The Modularity of Mind* (Cambridge, MA, The MIT Press/Bradford, 1983).
11. John Haugeland, "The Nature and Plausibility of Cognitivism", *The Behavioral and Brain Sciences*, 2 (1978), 215-225.
12. Adam Morton, *Frames of Mind* (Oxford, Oxford UP, 1980).
13. Stephen S. Stich, *From Common Sense Psychology to Cognitive Science* (Cambridge, MA, The MIT Press/Bradford, 1983).
14. K. V. Wilkes, "Functionalism, Psychology, and the Philosophy of Mind", *Philosophical Topics*, 12 (1981), 147-168.

De la colección de PAPELES JPG en <http://padron.entretemas.com.ve>



## El conocimiento tácito (Modularidad y subdoxasticidad)

Martin Davies\*

Hay dos tesis estrechamente vinculadas con la base de la obra realizada por Chomsky en lingüística y en filosofía. La primera es la de que los hablantes ordinarios de una lengua natural conocen tácitamente una gramática —un conjunto de reglas o de principios— de su lengua. La segunda tesis es la de que esa gramática especificada es adquirida por un niño sobre la base de un aparato innato masivo, es decir, un conocimiento innato de una gramática universal. Este artículo versa sobre la primera tesis.

Con el correr de los años se han utilizado diversas expresiones para elaborar esta tesis. Lo que un hablante conoce tácitamente equivale a la competencia de ese hablante (1965, pág. 8); y en el contexto de una discusión sobre argumentos de filósofos presas de la confusión, Chomsky utiliza a veces la expresión *tener la cognición de (cognize)*. Los hablantes conocen hechos —en el sentido habitual de este término— y también tienen la cognición de los hechos relativos, por ejemplo, a lo que significan diversas oraciones completas. Además, tienen la cognición (aunque no conozcan en el sentido habitual del término) de los hechos que se siguen de aquellos hechos del primer tipo (1980, págs. 69-70).

Podemos considerar los hechos del primer tipo como lo que enuncian los teoremas de una teoría sistemática. La idea fundamental sería pues la siguiente. Los hablantes ordinarios conocen hechos (y tienen la cognición de hechos) que enuncian los teoremas. También tienen la cognición —aunque no los conozcan en el sentido usual— de los hechos que establecen los

\* Publicado en la Festschrift en honor de Noam Chomsky. En la edición original francesa, este artículo fue traducido del inglés por P. Engel.

axiomas de los cuales derivan esos teoremas dentro de una teoría sistemática.

Si pensamos el problema en estos términos, es fácil atacar la primera de las dos cuestiones principales ante las que se encuentra enfrentado todo teórico que trabaja en este dominio. Considerando que habrá diversos conjuntos de axiomas de los que se pueden derivar los mismos teoremas sobre, por ejemplo, la significación de las oraciones completas, ¿hay una justificación empírica para suponer que un hablante ordinario sabe o tiene la cognición o —para decirlo en un vocabulario más neutro— ha interiorizado un conjunto de axiomas y que partiendo de aquél puede derivar los mismos teoremas pertinentes?

Ese es esencialmente el *desafío de Quine* (Quine 1972) en procura del concepto de conocimiento tácito. Mi objeto es intentar recoger ese desafío y tratar la noción de conocimiento tácito como cierta clase de estructura causal y explicativa subyacente o anterior a los conocimientos que un hablante pueda tener sobre oraciones completas. Ciertamente hay razones empíricas para atribuir a un hablante una determinada estructura antes que otra (veáanse Evans 1981, Davies 1987).

Dentro del marco del presente artículo, me contentaré con suponer que se puede dar este género de respuesta al desafío de Quine. Supondré también que se puede responder de manera satisfactoria a lo que cabe llamar *preocupaciones wittgensteinianas* (véase, por ejemplo, Hacker y Baker, 1984), que sugieren que es posible mostrar que la noción misma de conocimiento tácito es el producto de una confusión conceptual. Y supondré asimismo que es posible analizar el papel de las descripciones de conocimientos tácitos en las explicaciones psicológicas, especialmente que se puede especificar cómo la descripción del conocimiento tácito se vincula con otros niveles de descripción y de explicación, como los tres niveles distinguidos por Marr (1982).

Si, en consecuencia, se supone que las atribuciones de conocimiento tácito tienen una índole empírica respetable, que no son conceptualmente confusas y que desempeñan un papel legítimo y especificable dentro de una teoría psicológica, podemos abordar la segunda gran cuestión, que es la de saber cómo describir del mejor modo posible la relación psicológica entre una persona —un sujeto pensante, un hablante— y los axiomas

o reglas de la gramática interiorizados por esa persona. Los estados de conocimiento tácito son estados que poseen un contenido semántico. Pero ¿hasta qué punto se parecen a otros estados provistos de contenido o hasta qué punto difieren de ellos? Me refiero a estados como actitudes proposicionales, creencias, deseos, esperanzas, anhelos, etc.

### 1

La respuesta del propio Chomsky a esta cuestión no es simple. Uno de los componentes de la posición de Chomsky es la doctrina de que aquello que es objeto de la cognición o es tácitamente conocido es algo que implica creencias ordinarias. De conformidad con esta doctrina, Chomsky declara: (1976, pág. 163):

“Cabe concebir que creencias conscientes formen una subparte desperdigada (y probablemente sin interés) de la estructura cognitiva total”.

Y con respecto a una posible restricción del término “conocer” en el conocimiento consciente Chomsky observa (1976, pág. 165):

“En ese sentido, lo que es ‘conocido’ será una subparte mal definida y tal vez dispersa y caótica de los sistemas coherentes e importantes que forman la cognición. Para la psicología, la noción importante será ‘tener la cognición de’ y no ‘conocer’”.

Prevalece aquí la imagen de una estructura cognitiva compleja, algunas de cuyas partes afloran a la superficie en el nivel de la conciencia como creencias ordinarias. La distinción entre los estados que son conscientes y aquellos que no lo son probablemente no sea importante para la filosofía, y ciertamente esa distinción no es importante para la psicología.

Otros componentes del análisis incluyen la enunciación explícita (1980, pág. 93) de que “no ‘creemos en’ las reglas de gramática y los otros sistemas cognitivos”, como por ejemplo, la tajante distinción (1986, pág. 266) entre el hecho de conocer la regla R y el hecho de saber que la regla R existe o vale o que es una regla de nuestro lenguaje.

En una reciente presentación, esta distinción se relaciona con la imagen de una estructura deductiva compleja (1986, pág. 269):

“Se ha sostenido que es falso o hasta ‘ridículo’ decir que una persona conoce las reglas de gramática, aun en el sentido de un conocimiento implícito o tácito. Como enunciación de alcance general, esto no puede ser correcto. No vacilamos en decir que Juan sabe, cuando Pedro no sabe, que los verbos no pueden separarse de sus objetos por adverbios o que las ligaduras son aspiradas salvo después de *s*, *suponiendo*, *por supuesto*, *que conocemos la significación de los términos utilizados en esas atribuciones de conocimientos*. Observemos bien que sería falso decir que Juan sabe que la regla vale, pero ésta es otra cuestión. El hecho de saber si es igualmente apropiado utilizar el término ‘conocimiento’ o el término ‘saber’ en otros casos diferentes de los discutidos me parece poco claro porque el concepto es poco claro, aunque en todo caso de escasa importancia” (la bastardilla es mía).

Este pasaje indica y descarta varias supuestas fuentes de vacilación filosófica tocante a las atribuciones de conocimiento tácito. Una fuente que debemos considerar es la de que ponemos reservas respecto de la atribución de conocimiento tácito en el caso en que nosotros mismos no comprendamos los términos utilizados en la atribución.

Chomsky rechaza esta posible fuente de vacilación alegando que ella no tiene nada que ver en particular con el conocimiento del lenguaje. Sobre este punto ciertamente Chomsky tiene razón. En efecto, quienquiera que sostenga algo necesita saber la significación de los términos que emplea, pues de otra manera no sabe lo que quiere decir. Así, un hablante puede no atribuir el conocimiento de reglas lingüísticas particulares a otro hablante alegando sencillamente que el otro hablante no comprende los términos empleados en la formulación de esas reglas. Pero esto nada tiene que ver con la tesis de que las atribuciones de conocimiento tácito son semejantes a las atribuciones de actitudes proposicionales ordinarias.

Chomsky tiene pues entera razón cuando dice que esta fuente posible de vacilación filosófica carece de fundamento. Pero hay un punto que en cambio resulta muy pertinente. La cuestión importante no es saber si la persona que hace la atribución necesita comprender los términos utilizados en ella, es decir, si tiene necesidad de comprender los conceptos entrañados en los contenidos de los estados que son el objeto de la atribución. Lo que importa es saber si la persona a quien se atribuye el estado en cuestión tiene necesidad de comprender esos conceptos.

Corrientemente se afirma que nadie puede tener una creencia provista de un determinado contenido —o tener un pensamiento con cierto contenido— sin captar los conceptos que constituyen dicho contenido. Pero si es posible atribuirles a los hablantes usuales un conocimiento tácito de reglas, principios o generalizaciones lingüísticas, no se puede exigir esto en lo referente a contenidos de estados de conocimiento tácito, pues la mayor parte de los hablantes corrientes no tienen ningún dominio de los conceptos propios de la teoría lingüística.

El requisito de que la persona que piensa y habla debe conceptualizar lo que piensa y dice es la señal de una diferencia intuitiva entre los estados ordinarios de actitudes proposicionales y los estados de conocimiento solamente tácito. Esta es una diferencia de la que luego volveré a ocuparme. Pero intentemos primero bosquejar un marco dentro del cual podamos establecer la diferencia.

Necesitamos investigar algunas de las diferencias que hay entre los estados de actitudes proposicionales ordinarias (como las creencias) y otros estados psicológicos cognitivos, como los estados de conocimientos tácitos de reglas lingüísticas. El objeto de esta investigación es descubrir si puede haber una distinción fundada —en el plano filosófico o en el psicológico— entre los estados *subdoxásticos*.

Esta investigación debe suponer que tenemos ciertos ejemplos de lo que a primera vista parecen ser estados subdoxásticos, es decir, estados cognitivos psicológicos que no sean estados de actitudes proposicionales. Pero la investigación no debe partir de peticiones de principio sobre lo que entra o no dentro del dominio del esquema de atribuciones de actitudes del sentido común o de la explicación racional.

El análisis del concepto de conocimiento tácito que ofrece Evans (1981) se basa en un examen de las teorías semánticas como ejemplo central. La ventaja de este ejemplo consiste en que la estructura de dichas teorías es familiar a los filósofos. Por lo demás, desde un punto de vista intuitivo sería problemático atribuir a los hablantes ordinarios actitudes proposicionales cuyos contenidos estén dados por los axiomas de una teoría semántica. Frente a esas intuiciones, el análisis del conocimiento tácito de Evans permite atribuir a hablantes el conocimiento tácito de una determinada teoría antes que de otra en ciertas condiciones empíricas.

Sin embargo se puede sostener que ese análisis podría aplicarse a alguien que tuviera un conocimiento consciente de reglas semánticas. Lo que este análisis nos suministra en realidad son condiciones de algo que pudiera ser *por lo menos* conocimiento tácito. Por consiguiente, no se desprende del análisis en cuestión que exista una distinción fundada entre estados de actitudes y estados subdoxásticos, en tanto que el conocimiento de las reglas o generalizaciones semánticas entra dentro de lo subdoxástico, es decir, de aquello que es *simplemente* tácito. Se ha podido sostener (Campbell 1982) que, a pesar de las intuiciones iniciales, el conocimiento de las reglas semánticas forma parte del dominio de las actitudes proposicionales. De modo que la investigación sobre una división fundada no debe partir de supuestos referentes a la localización de nuestro conocimiento de las reglas semánticas, es decir, si ese conocimiento cae en una o en otra parte de esta división.

## 2

Hay por lo menos tres diferencias entre estados de actitudes proposicionales y ejemplos *prima facie* de estados subdoxásticos. La primera diferencia es muy evidente.

Los estados psicológicos difieren en relación con su *accesibilidad a la conciencia*. Las creencias corrientes son estados conscientes; en cambio una parte de aquello de que se tiene cognición “queda siempre oculto a la conciencia” (Chomsky 1976, pág. 165). La idea de que la conciencia debe marcar la frontera entre el dominio de las actitudes y el dominio simplemente cognitivo es tentadora. Pero esta idea es también problemática, pues, como hemos de verlo seguidamente, las aguas que rodean la noción de accesibilidad a la conciencia pueden

fácilmente llegar a estar revueltas. Por último, es dudoso que esta idea tentadora pueda tener algún peso teórico. En este sentido, el escepticismo de Chomsky en cuanto a la importancia de una distinción entre el conocimiento consciente y el conocimiento tácito está justificado.

Hay una segunda diferencia entre los estados psicológicos, diferencia que se relaciona con su *integración inferencial* o con su promiscuidad, por oposición a su *aislamiento inferencial* (véase Stich 1978). Se puede ilustrar la idea de promiscuidad inferencial del modo siguiente. Un estado de creencia, según la expresión de Evans (1981), está “al servicio de una multiplicidad de proyectos distintos”. La creencia de que unos determinados objetos son zapatos forrados de piel y de color morado, puede combinarse con otra creencia (por ejemplo, la creencia de que los únicos ejemplares de zapatos morados forrados de piel que hay en los alrededores son las pantuflas de la abuela) y puede llevar inferencialmente a una nueva creencia, la de que ellos son las pantuflas de la abuela. Esta última creencia puede a su vez combinarse con un deseo (el de verter champaña en las pantuflas de la abuela), de suerte que se produzca una razón para obrar. La creencia inicial también podría combinarse con creencias muy diferentes —por ejemplo la de que esos objetos morados y forrados de piel pueden divertir al gato— para llevar a diferentes consecuencias inferenciales y para contribuir a dar razones para realizar diferentes acciones.

Las creencias forman una red, que posee conexiones inferenciales y justificadoras; ésta es la concepción habitual. Y una parte de esta concepción consiste en que una nueva creencia —por ejemplo la de que la abuela acaba de entrar con sus pantuflas puestas— puede conducir inferencialmente al ajuste de creencias sostenidas anteriormente, a saber, que esos objetos morados y forrados no son en definitiva realmente zapatos.

Las creencias se combinan para producir consecuencias inferenciales, y las creencias se ajustan unas a otras para evitar la incoherencia. Estos son dos aspectos de la integración inferencial de las creencias. Pero la idea de una creencia puesta al servicio de una multiplicidad de proyectos implica otra idea, que la noción de una red de creencias no traduce verdaderamente. En efecto, una creencia puede combinarse con varios deseos diferentes de manera que se formen razones para realizar numerosas acciones distintas. Esta tercera propiedad de las

creencias es una propiedad notable del dominio de las actitudes, y es ese rasgo lo que impide directamente reducir enunciaciones sobre creencias a enunciaciones sobre la disposición a una determinada conducta.

Los estados psicológicos cognitivos que son en primer término estados de simple cognición no parecen poseer estas propiedades características de las creencias. Por ejemplo, estados de conocimiento tácito sobre reglas o principios sintácticos o estados propios de los estadios iniciales del tratamiento de la información visual no se combinan directamente con deseos para producir razones de obrar. Lo que contribuye a la formación de esas razones son sólo las consecuencias remotas de tales estados. Estos ejemplos de lo que a primera vista pueden considerarse estados subdóxicos parecen pues carecer de la tercera propiedad de las creencias.

No es intelectualmente difícil tener una creencia incompatible con lo que se conoce tácitamente. Así, los estados de creencia tácita no parecen poseer la segunda propiedad de las creencias o, mejor dicho, no tienen esta propiedad cuando se trata de creencias ordinarias. Establecer si los estados de conocimiento tácito pueden tender a una coherencia interna sería otra cuestión difícil de resolver mediante ejemplos.

Los estados de conocimiento tácito se combinan ciertamente con otros para producir consecuencias. Esto forma parte de la imagen compleja que nos ofrece Chomsky de los estados de la cognición. Por ejemplo, no podemos decir sencillamente que los estados subdóxicos no posean la primera propiedad de las creencias. Podemos observar que las combinaciones que se producen dentro de los estados subdóxicos parecen simples en lugar de estar sujetas a la promiscuidad, por más que se pueda replicar que éste es simplemente el resultado de un pequeño número de estados disponibles para semejantes combinaciones. Pero lo que importa no son esos lazos inferenciales que existen en el interior de los estados de conocimiento subdóxico. Lo más notable es que los estados de conocimiento tácito no parecen estar inferencialmente integrados en la red de las creencias. Antes bien, parecen aislados de esa red. Y ese aislamiento inferencial tiene dos aspectos.

Por un lado, los estados subdóxicos no se combinan inferencialmente con creencias para producir otras creencias. Por ejemplo, nuestra identificación de palabras pronunciadas

implica en cierto estadio el registro de una información de propiedades acústicas, tales como la cantidad de emisión de voz. Por ejemplo, una persona que oye una palabra que comienza con una vocal elidida registrará la información de cuál es la emisión de la voz, digamos, 45 milisegundos. Esa persona puede, por cualquier razón, creer que si la sílaba que sigue corresponde a una emisión de 45 milisegundos es la hora de marcharse para ir a almorzar. Pero el hecho solo de oír el fonema elidido como tal no bastará para dar a la persona en cuestión una razón para dejar sus herramientas de trabajo.

La idea que está en la base de este ejemplo es la de que si un sujeto tiene dos creencias —una de la forma *que si P entonces Q* y la otra de la forma *que P, — entonces el sujeto inferirá a menudo la conclusión de que Q*. Pero si las dos premisas del *modus ponens* están divididas (si una es el contenido de una creencia y la otra es el contenido de un estado subdoxástico), entonces el sujeto no llega a creer en la conclusión.

Por otro lado, las creencias ordinarias no se combinan habitualmente —o por lo menos frecuentemente— con estados subdoxásticos para producir otros estados subdoxásticos. En realidad, los procedimientos que operan en los estados que representan en primer término ejemplos de estados subdoxásticos son independientes de lo que pueda creer una persona.

La tercera diferencia dentro de la clase de los estados psicológicos es aquella a que ya aludimos referente al *dominio de los conceptos*. Un hecho fundamental y característico de las actitudes proposicionales es que ningún sujeto pensante puede tener una actitud ante una proposición si no tiene un dominio mínimo de los conceptos constitutivos de esa proposición. Un pensador que no disponga del concepto de verde sencillamente no puede creer que la hierba sea verde. Puede por supuesto creer que la oración “la hierba es verde” expresa una proposición verdadera o no; pero no puede saber de qué proposición se trata. Asimismo, un sujeto que no disponga del concepto de trinidad —que no sabe lo que sea la trinidad de una determinada cosa— no puede creer que Dios sea una trinidad. Puede creer que la oración “Dios es trinidad” expresa otra proposición, pero sin saber de qué proposición se trata.

En cambio, se pueden atribuir a un sujeto estados de conocimiento tácito cuyos contenidos implican conceptos que el

sujeto ignora por entero. La mayor parte de los conceptos empleados en la construcción de reglas y de principios lingüísticos son conceptos que la gran mayoría de los hablantes corrientes no poseen.

Hay pues tres diferencias intuitivas entre los estados de actitud —en particular las creencias— y los ejemplos de lo que se puede considerar en primer término estados subdoxásticos. Las creencias son accesibles a la conciencia, mientras que los estados subdoxásticos no lo son. Las creencias están inferencialmente integradas, mientras que los estados subdoxásticos están inferencialmente aislados de las creencias. Tener una creencia exige dominio de los conceptos constitutivos de esa creencia, mientras que estar en un estado subdoxástico no exige dominar los conceptos entrañados en su contenido.

### 3

Ya hemos observado que la noción de accesibilidad a la conciencia es delicada. Podemos realmente preguntarnos si el contraste entre accesibilidad e inaccesibilidad a la conciencia es la marca de una diferencia teórica importante. Veamos tres razones por las cuales resulta poco sensato asignar un peso teórico cualquiera a esa noción.

La primera razón es sencillamente el hecho de que no empleamos —o sólo los psicoanalistas la emplean— la noción de creencia inconsciente. Pero esos estados son intuitivamente por completo diferentes de los estados psicológicos que pudiéramos tomar como ejemplos de subdoxasticidad. Por consiguiente, la inaccesibilidad a la conciencia no basta por sí sola como señal de la subdoxasticidad.

La segunda razón consiste en que hay rasgos de la red inferencial de las creencias que son típicamente inconscientes. Fodor da como ejemplo (1983, pág. 85) nuestra “aquiescencia inconsciente a la regla del *modus ponens*”. La idea subyacente es la siguiente. La regla del *modus ponens* es algo de lo que el hablante ordinario no tiene conciencia. Un sujeto puede hacer sistemáticamente inferencias que son en realidad instancias del *modus ponens* y puede ofrecer las premisas de una determinada inferencia como razones para creer en su conclusión sin ser capaz de expresar la regla en su generalidad. En este sentido, la regla inferencial es algo de lo que la persona corriente no tiene conciencia; desde este punto de vista, dicha regla es

semejante a una regla lingüística. Pero, a pesar de esto, nuestro respeto de la regla del *modus ponens* es intuitivamente y por entero diferente de los anteriores ejemplos de subdoxasticidad.

Se puede dar una respuesta muy natural a esta segunda razón para dudar de que la noción de inaccesibilidad a la conciencia pueda tener un peso teórico cualquiera. Podemos observar que, en psicología del tratamiento de la información, existe una distinción entre, por un lado, los estados psicológicos provistos de contenido o representacionales, y por otro, los procesos o cálculos (*computations*) sobre esos estados. En muchos aspectos la distinción es análoga a la distinción, en lógica, entre una fórmula y una regla de inferencia. Y como el *modus ponens* es una regla de inferencia, el hecho de que la observemos debe considerarse como un hecho que revela la presencia de un procesador calculatorio antes que un estado representacional.

La idea que está en la base de esta primera respuesta es que la distinción entre creencias y estados subdoxásticos es, en primer lugar, una distinción entre estados de creencia —sobre los cuales se efectúan inferencias— y otros estados representacionales, sobre los cuales se efectúan cálculos. Por consiguiente, la inaccesibilidad a la conciencia debe considerarse en primer lugar como señal de la subdoxasticidad de los estados representacionales. La inaccesibilidad a la conciencia de una regla inferencial no amenaza la condición de esta regla como propiedad de la red de las creencias.

Lo que hace resaltar esta primera respuesta es importante. Pero esta respuesta no equivale a asignar un peso teórico cualquiera a la noción de accesibilidad a la conciencia. Pues si la inaccesibilidad a la conciencia es solamente una marca de subdoxasticidad en el caso de los estados representacionales, luego se sigue que la inaccesibilidad a la conciencia de las reglas lingüísticas no es pertinente ella misma. En efecto, no hay ninguna razón para suponer que (en un modelo de tratamiento de la información) una regla lingüística interiorizada se manifieste como un estado representacional antes que como un procesador calculatorio. De manera que nuestra primera respuesta no hace sino rebajar un poco la importancia de la accesibilidad a la conciencia.

Se podría tratar de defender aún esta noción agregando lo siguiente. Si la inaccesibilidad a la conciencia nos da por lo

menos una distinción clara entre las creencias y los otros estados representacionales, dicha inaccesibilidad nos ofrece también una distinción (derivada) entre reglas de inferencias mentales auténticas y procesos calculatorios solamente cognitivos. Pues, atendiendo a la distinción inicial, resultará fácil clasificar una regla o un proceso según que su dominio sea una clase de creencias, por un lado, o una clase que contiene estados representacionales subdoxásticos, por otro.

Pero, por supuesto, la significación de este segundo punto depende de la cuestión de establecer si la accesibilidad a la conciencia permite distinguir claramente unos estados representacionales de otros. Y la tercera razón por la que es poco razonable asignar a esa noción un peso teórico cualquiera se refiere directamente a los estados representacionales.

Fodor hace notar (1983, pág. 58) que hay ejemplos que hacen plausible la siguiente idea:

“El hecho de que los niveles inferiores de análisis periférico (*input analysis*) sean relativamente inaccesibles depende por lo menos en parte de las prioridades que se privilegien en la transferencia de las representaciones desde la memoria de plazo (relativamente) corto a la memoria de plazo (relativamente) largo”.

Ahora bien, Fodor sostiene que desde el punto de vista de la teoría psicológica hay razones más profundas de esta relativa inaccesibilidad. Pero a los efectos de nuestro razonamiento presente sólo tenemos necesidad de observar un hecho, a saber, que las prioridades en la transferencia de las representaciones desde la memoria de corto plazo a la memoria de largo plazo son contingentes. Y bien podemos imaginar que esas prioridades sean diferentes.

Supongamos, en consecuencia, que en efecto esas prioridades son diferentes. Supongamos que los estados representacionales de nivel inferior sean transferidos a la memoria de relativamente largo plazo y por consiguiente afloran en la conciencia, acaso como raspaduras, como rasguños o comezones reconocibles. Ese cambio empírico sin duda no sería suficiente para transformar tales estados en creencias. Por consiguiente la inaccesibilidad a la conciencia de un estado representacional no es en sí misma decisiva.

Alguien podría replicar que, en el caso imaginado, los estados conscientes —como son sólo rasguños o comezones— no reflejan en la conciencia el contenido semántico que se supone tenían originariamente los estados inaccesibles. Pero no es seguro que esta objeción esté justificada. En efecto, supongamos que el carácter fenomenológico de los estados conscientes dependa sistemáticamente del contenido semántico del estado inicial en un nivel inferior. Esto no hace del estado consciente una creencia, pues hay una diferencia radical entre el hecho de que cierto contenido semántico se refleje en la experiencia de un sujeto y el hecho de que ese sujeto conceptualice dicho contenido. Sin embargo esta circunstancia sirve sencillamente para sugerir que, aun en el caso de los estados representacionales, lo que cuenta no es la accesibilidad a la conciencia como tal. Lo importante parece ser antes bien la diferencia entre los estados cuyos contenidos están conceptualizados por el sujeto y los estados que no lo están.

Por estas tres razones debemos rechazar el prejuicio de que la noción de accesibilidad a la conciencia pueda servirnos para fundar una diferencia entre creencias y estados subdoxásticos. Si hacemos a un lado el concepto de accesibilidad a la conciencia, quedan dos diferencias intuitivas entre creencias y estados subdoxásticos. Las creencias están inferencialmente integradas unas con otras, mientras que los estados subdoxásticos están inferencialmente aislados de las creencias. Y tener una creencia requiere un dominio de los conceptos componentes mientras que el contenido de un estado subdoxástico no está conceptualizado por la persona en cuestión. Estas dos distinciones sugieren otras dos líneas de investigación bastante diferentes. La idea de conceptualización sugiere un análisis más a priori o constitutivo de la diferencia entre las creencias y los estados subdoxásticos. La idea de integración inferencial sugiere una manera de remitirse a la ciencia empírica.

#### 4

Evans vincula la noción de conceptualización con lo que él llama la *condición de generalidad* (*Generality Constraint*). De manera que, según Evans, la idea de que una creencia es un estado en el que el sujeto conceptualiza su contenido implica en parte lo siguiente (1981, pág. 133):

“No estaremos dispuestos a atribuir a un sujeto la creencia de que *a es F* (en el caso de un objeto particular *a* y en el caso de una propiedad *F*) a menos que podamos suponer que el sujeto es capaz de sustentar la suposición (o tener el pensamiento de) de que *b es F* en el caso de todo objeto del cual el sujeto tenga un concepto cualquiera”.

La idea familiar expuesta aquí es que el hecho de tener una creencia —o en general de sustentar un pensamiento— implica el despliegue de conceptos. Para tener el pensamiento de que *a es F* un sujeto debe saber lo que es para un objeto ser *F*, es decir, lo que es para un objeto *arbitrario* ser *F*. Pero si un sujeto sabe lo que es para un objeto ser *F* y si es capaz de pensarlo con referencia al objeto *b*, luego ese sujeto es capaz de sustentar el pensamiento de que *b es F*.

Esta idea fundamental sobre la conceptualización está desarrollada en *The Varieties of Reference* (1982), donde se presenta explícitamente la condición de generalidad como condición que se refiere a los contenidos de los estados de actitud (pág. 104):

“Si se puede atribuir a un sujeto el pensamiento de que *a es F*, ese sujeto debe poseer los recursos conceptuales para tener el pensamiento de que *a es G* en el caso de toda propiedad *G* de que el sujeto tenga un concepto. Vemos, pues, que el pensamiento de que *a es F* se encuentra en la intersección de dos series de pensamientos: por un lado, la serie de pensamientos de que *a es F*, de que *b es F*, de que *c es F*... y, por otro lado, la serie de pensamientos de que *a es F*, de que *a es G*, de que *a es H*...”.

Se puede expresar esta doctrina afirmando que los pensamientos están esencialmente estructurados; los pensamientos tienen elementos constitutivos que pueden recombinarse para producir nuevos pensamientos. De manera que una consecuencia inmediata de la condición de generalidad es la de que, si se puede atribuir a un sujeto el pensamiento de que *a es F* y el pensamiento de que *b es G*, luego ese sujeto posee los recursos conceptuales suficientes para sustentar el pensamiento de que *a es G* y el pensamiento de que *b es F*.

Seguimos aquí una línea de indagación sugerida por una de las dos diferencias intuitivas entre las creencias y los estados subdoxásticos, a saber, el contenido de una creencia está conceptualizado por la persona en cuestión, en tanto que el contenido de un estado subdoxástico no lo está. La introducción de la condición de generalidad conduce a la idea de que dicha condición suministra una distinción fundada entre las dos esferas. Y ocurre que el propio Evans sugirió que la condición de generalidad señala una distinción entre los contenidos de los estados de actitud proposicional y los contenidos de los estados entrañados en “el tratamiento de información que se realiza en nuestros cerebros” (1986, pág. 104, nota 22; véase Davies 1986, págs. 143-6).

El punto central aquí es que queda abierta la siguiente posibilidad. Un sujeto puede ser capaz de tener el pensamiento de que  $a$  es  $F$  y tal vez en un estado subdoxástico que tenga el contenido de que  $b$  es  $G$ . Sin embargo puede ocurrir que el sujeto tenga el pensamiento de que  $a$  es  $G$  porque no dispone del concepto *de ser un G*; y puede ocurrir que no haya un estado subdoxástico real o posible cuyo contenido sea  $a$  es  $G$ , porque el sistema que trata la información de que  $b$  es  $G$  no contenga sencillamente información sobre el objeto  $a$ .

Este punto central ilustra la manera en que el dominio total de los estados de actitudes proposicionales y de los estados subdoxásticos puede violar la condición de generalidad. Esto ocurre sólo si nos concentramos exclusivamente en un sistema único de tratamiento de la información y entonces descubriremos casos reales en que no se respeta la condición de generalidad. En cambio, es posible imaginar sistemas que obedecen a las siguientes condiciones. Si hay estados reales o potenciales del sistema, estados que tienen los contenidos de que  $a$  es  $F$  y de que  $b$  es  $G$ , luego hay otros estados reales o potenciales del sistema que poseen los contenidos de que  $a$  es  $G$  y de que  $b$  es  $F$ . De esta posibilidad ¿se desprende que la condición de generalidad no nos permite diferenciar los estados de actitud de los estados subdoxásticos?

La respuesta depende de la manera en que se suponga que la condición de generalidad traduce esta distinción. Supongamos que se sostenga que la diferencia vital entre estados de actitudes y estados subdoxásticos se reduzca a lo siguiente: que los contenidos de estados de actitud satisfacen el requisito de

cierre que implica la condición mientras que los contenidos de estados subdoxásticos no lo satisfacen. En este caso, el tipo de ejemplo antes mencionado invalidaría la idea de que la condición de generalidad traduce una diferencia importante. Pero una defensa de la condición de generalidad debe ser más compleja.

En el sistema imaginado, los contenidos de los estados satisfacen en realidad el requisito de cerrado. Pero queda la circunstancia de que asignar contenidos a esos estados no depende de ese hecho. Se puede decir que satisfacer ese requisito de cierre es un rasgo intrínseco de esos estados provistos de contenido. En cambio, asignar un contenido semántico a un estado de actitud es algo que procede directamente de la condición de generalidad. Creer que  $a$  es  $F$  es ejercer *ipso facto* dominio del concepto *ser F*; y el sujeto puede ejercer ese dominio conceptual en estados de actitud.

En consecuencia, emerge aquí la idea de que una distinción fundada entre estados de actitud y estados subdoxásticos debe basarse en una distinción entre dos tipos de contenidos. La condición de generalidad gobierna en parte la noción de contenido conceptualizado que se aplica a los estados de actitud. Pero su relación con el concepto de contenido en el caso de los estados de actitud es sólo accidental.

Esta idea nos lleva a considerar la tercera de las diferencias intuitivas entre las creencias y los ejemplos anteriores de estados subdoxásticos. Se puede considerar que la condición de generalidad hace explícito lo que implica el dominio de los conceptos o la idea de conceptualización. Pero de ello no se sigue que esa distinción confirme de manera decisiva nuestros juicios intuitivos sobre la relación entre el conocimiento tácito de las reglas sintácticas y la subdoxasticidad. Pues progresivamente se va destacando una distinción entre, por una parte, los estados representacionales y, por otra, los procesos o procedimientos calculatorios en los que se basan esos estados. La distinción entre lo que está conceptualizado y lo que no lo está es ante todo una distinción entre estados. Y sin embargo, el conocimiento de reglas sintácticas podría darse en virtud de la presencia de un procesador antes que por la presencia de una colección de estados. De manera que el hecho de que los contenidos de conocimientos tácitos no estén conceptualizados por la persona en cuestión no muestra de manera decisiva que



esos conocimientos tácitos pertenezcan a una determinada categoría o a otra.

La clasificación de ciertos conocimientos tácitos es, pues, un problema secundario que depende de una clasificación preliminar de los estados que desempeñan el papel de entradas y de salidas de los procesadores que realizan el conocimiento tácito. En el caso de un conocimiento tácito de reglas y de principios lingüísticos, la clasificación preliminar dependerá a su vez directamente de la cuestión de establecer si las descripciones de las oraciones, en virtud de las cuales las oraciones forman parte de generalizaciones lingüísticas, son descripciones que la persona conceptualiza. Sin duda alguna, el resultado será el de que muchos conocimientos tácitos entrarán en la categoría de los estados subdoxásticos.

La cuestión correspondiente en el caso de la regla del *modus ponens* es la de saber si las descripciones del mundo que desempeñan el papel de premisas para las inferencias sometidas a la regla están ellas mismas conceptualizadas por la persona en cuestión. Como la respuesta a esta pregunta es afirmativa, el conocimiento que tiene un hablante de esta regla entrará en la categoría de los estados de actitud, antes que en la categoría de los estados subdoxásticos; la regla opera más en la primera esfera que en la segunda.

## 5

El contraste entre la integración inferencial y el aislamiento inferencial sugiere una línea de análisis diferente. En la sección 2, hemos señalado dos componentes del aislamiento inferencial en ejemplos de estados subdoxásticos comparados con estados de creencia. Por un lado, los estados subdoxásticos no se combinan inferencialmente con creencias para producir nuevas creencias. Por otro lado, las creencias ordinarias no se combinan en general con estados subdoxásticos para producir otros estados subdoxásticos. Este segundo componente nos recuerda el concepto de encapsulamiento informacional de Fodor: un sistema que opera en el interior de un organismo está informacionalmente encapsulado si la información de que dispone el sistema es considerablemente menor que la totalidad de la información representada en el organismo (Fodor 1983, página 69).

Para Fodor, el encapsulamiento informacional es una de las propiedades que definen la *modularidad*. En realidad, es “lo que yo considero como acaso el aspecto más importante de la modularidad” (pág. 37) y “la esencia misma de la modularidad” (pág. 71). En consecuencia, la esencia de un módulo es un sistema computacionalmente articulado (y por lo tanto, más complejo que un simple reflejo) e informacionalmente encapsulado (*encapsulated*), una propiedad que comparte con los reflejos.

La división principal dentro de la arquitectura esquemática del espíritu de Fodor es la división entre el sistema cognitivo central y los sistemas periféricos (*input systems*). Un sistema periférico es un sistema de tratamiento de la información que entrega en última instancia información al sistema cognitivo central, cuya función principal es a su vez fijar la creencia. En este esquema general, el sistema cognitivo central parece corresponder al ámbito de las actitudes proposicionales y, por supuesto, Fodor atribuye al sistema central las propiedades típicas del dominio de las actitudes, en particular la propiedad del holismo en relación con los hechos o la propiedad de la isotropía (pág. 105). Esto limita de manera sugestiva los sistemas periféricos a la esfera de los estados subdoxásticos.

Ahora bien, la tesis principal de *la modularidad del espíritu* postula que los sistemas periféricos son módulos. Esto equivale a afirmar, entre otras cosas, que los sistemas periféricos tienen un dominio específico, que sus operaciones son rápidas y automáticas, que sus niveles intermedios son relativamente inaccesibles al sistema central y tienen un ritmo, un desarrollo en secuencias y esquemas de ruptura característicos. Todo esto distingue la modularidad. Pero el punto más importante es el de que los sistemas periféricos estén informacionalmente encapsulados.

El concepto de modularidad de Fodor no es ciertamente el único concepto de modularidad que se haya empleado en las teorías psicológicas. Pero supongamos que se trate de una noción que tiene algún peso en las teorías y las explicaciones psicológicas. Si aventuramos esta hipótesis es tentador sugerir que la categoría científica y psicológica de modularidad es aquello que está en la base del concepto filosófico de subdoxasticidad. En efecto, nuestros ejemplos iniciales de estados subdoxásticos se refieren a estados de los sistemas periféricos:

los sistemas perceptuales más el lenguaje (Fodor 1983, pág. 44). Y de conformidad con la tesis de la modularidad, los sistemas periféricos tienen la propiedad del encapsulamiento informacional, que recuerda uno de los componentes del aislamiento inferencial que es una de las marcas de la subdoxasticidad.

Podemos seguir así la sugestión de que la modularidad está en la base de la subdoxasticidad. Hemos señalado ciertas características filosóficas intuitivas de la distinción entre creencias y estados subdoxásticos, entre la esfera de las actitudes y la esfera del simple conocimiento tácito. Pero por lo menos una de estas características no tendrá en definitiva gran peso teórico. Se plantea entonces la cuestión de si esas características intuitivas pueden darnos respuestas decisivas cuando tenemos que vérnoslas con casos límites, como por ejemplo nuestros conocimientos de las reglas semánticas. Cuando reflexionamos como filósofos sobre la subdoxasticidad somos algún tanto parecidos al hombre del sentido común de Locke, que tiene una idea del oro según los rasgos de conjunto observables de este metal.

Verdad es que la teoría de Locke sostiene que cada conjunto de rasgos observables globales —cada esencia nominal— está subtendida por una disposición cualquiera de partículas invisibles (una microestructura o una esencia real) pero que, en el nivel de la microestructura, la naturaleza es infinitamente variable (Ayers, 1981). Según esta teoría, todo esquema de clasificación en el nivel de la microestructura es una proyección partiendo del nivel de los rasgos globales observables.

Pero los filósofos, tanto anteriores como posteriores a Locke, han aceptado la idea de que pueda haber clasificaciones o agrupaciones en la naturaleza, la idea de que hay especies naturales. Esto hace posible una teoría particular de la semántica de vocablos tales como “oro”. De acuerdo con esta teoría el término “oro” designa la sustancia natural química que tiene un determinado número atómico. El hombre del sentido común no tiene necesidad de saber nada de química; tan sólo oye designar con el término “oro” la sustancia química que en los casos habituales ejemplifica un conjunto particular de rasgos globales observables. En términos de esta especie natural, los expertos pueden explicar por qué esos rasgos globales observables están de acuerdo entre sí, cuando lo están. Y cuando se las tienen que ver con casos límites de materia, casos que en su mayor parte,

pero no todos, poseen estos rasgos observables, apelan al concepto de límite de una especie natural.

Se puede hacer una comparación entre la teoría de Locke sobre el oro y el concepto de subdoxasticidad del modo siguiente. Para toda combinación de rasgos psicológicos intuitivos —como la inaccesibilidad a la conciencia, el aislamiento inferencial en comparación con la red de las creencias o la falta de conceptualización— hay una explicación en el nivel de los procesos psicológicos subyacentes. Pero los procesos cognitivos varían de todas las maneras posibles, tanto que cualquier taxonomía que se sitúe en ese nivel subyacente debería proyectarse partiendo del nivel más intuitivo del sentido común.

En cambio, el análogo de la teoría antigua y de la teoría más reciente de que existen especies naturales podría ser el siguiente. En la naturaleza hay muchas especies de procesos psicológicos y esta clasificación desempeña su parte en las teorías psicológicas explicativas. En particular, esta taxonomía científica nos permite explicar por qué diversos rasgos intuitivos pueden agruparse, si es que pueden estarlo. Y esto nos permite también hacer una clasificación autorizada de los diversos casos límites.

El hombre de la calle que presume que existen especies químicas no tiene necesidad, como tal, de conocer la química. Del mismo modo, el filósofo ordinario no tiene necesidad de ser un experto en psicología cognitiva. Puede tan sólo suponer que hay grupos naturales dentro de la realidad psicológica y puede considerar la subdoxasticidad como una de esas especies naturales —sea lo que ésta fuere en última instancia— que, en los casos habituales, ejemplifica un conjunto particular de rasgos psicológicos ordinarios: inaccesibilidad a la conciencia, aislamiento inferencial en relación con la red de creencias y contenido no conceptualizado.

Para que esta sugestión tentadora —la de que la modularidad está en la base de la subdoxasticidad— sea correcta, sería menester que la modularidad fuera una especie natural en psicología y que fuera la esencia real cognitiva correspondiente a la esencia nominal considerada por el filósofo. El concepto psicológico cognitivo de modularidad daría así una explicación de las razones por las cuales los rasgos habituales de la subdoxasticidad se manifiestan juntos, si así lo hacen. Y esto nos daría una clasificación fundada (a posteriori y empírica) de los casos límites delicados.

Así se sugiere pues que la distinción entre el sistema cognitivo central y los sistemas periféricos modulares constituye la esencia real de la distinción entre creencias y estados subdoxásticos. Esto está corroborado por lo menos vagamente por la semejanza que hay entre el encapsulamiento informacional y el aislamiento inferencial, lo cual hace un llamado al espíritu interdisciplinario, pero por el momento éstas no son más que especulaciones.

La cuestión de establecer si esta sugerencia es realmente viable depende de varios factores. Es evidente que depende del poder explicativo del concepto de modularidad de Fodor dentro de una teoría psicológica. También depende de la posibilidad de dar explicaciones (desde el punto de vista de la modularidad) de los diversos rasgos de la subdoxasticidad. Y depende asimismo del éxito que alcance la noción de modularidad en el tratamiento de casos límites delicados.

No es este el lugar apropiado de un análisis general de la tesis de la modularidad. Pero si dejamos de lado la perspectiva de una explicación satisfactoria (desde el punto de vista de la modularidad) sobre los rasgos intuitivos de la subdoxasticidad, podemos tratar de utilizar la noción de modularidad para trazar un límite en este territorio fronterizo.

## 6

Es tentador remitirse a la ciencia, en parte porque una teoría científica puede dar respuestas definitivas sobre casos límites que son problemáticos a los ojos del filósofo medio. En el caso de nuestro conocimiento del lenguaje es plausible que nuestro conocimiento de reglas o de principios sintácticos quede clasificado entre los estados subdoxásticos. Y es plausible que nuestro conocimiento de principios pragmáticos corresponda a la esfera de las actitudes. El caso límite problemático es el de las reglas semánticas.

El análisis del conocimiento tácito de Evans (1981) está hecho partiendo del ejemplo central de las teorías semánticas. Y el propio Evans clasifica nuestro conocimiento de las reglas semánticas dentro de lo subdoxástico. Pues sobre las reglas semánticas y sobre las reglas sintácticas Evans decía lo siguiente (1981, págs. 338-339):

“El conocimiento tácito de las reglas semánticas y sintácticas del lenguaje no está constituido por estados

del mismo tipo que los estados que identificamos en nuestro uso ordinario con términos tales como ‘creencia’ y ‘conocimiento’. Poseer el conocimiento tácito es algo que se manifiesta exclusivamente por el hecho de hablar y de comprender una lengua; la información no está ni siquiera potencialmente al servicio de otro proyecto del agente, así como no puede entrar en interacción con otras creencias del agente para producir nuevas creencias. Conceptos como los que utilizamos al especificarlo no son conceptos de los que tengamos necesidad de suponer que el sujeto los posee, pues ese estado está inferencialmente aislado del resto de los pensamientos y de las creencias del sujeto”.

Aquí se tratan dos puntos. Uno se refiere a la falta de integración inferencial. El otro se refiere a la falta de conceptualización. La tesis de Evans sostiene que en estos dos registros el conocimiento de las reglas sintácticas y semánticas es sencillamente tácito.

Consideremos primero el punto referente a la conceptualización en el caso de las reglas semánticas. Estas reglas vinculan las palabras y las construcciones tales como figuran en las oraciones con los elementos constitutivos de los pensamientos expresados por las oraciones. Es claro que los elementos constitutivos de pensamientos están conceptualizados por la persona en cuestión. Si las reglas semánticas vinculan una expresión particular con la propiedad de ser feliz y si el hablante conoce tácitamente esa regla, esto no puede ocurrir si la persona no dispone del concepto de ser feliz.

Existen otros dos tipos de términos utilizados en las reglas semánticas. Los términos teóricos semánticos, como por ejemplo *denota*, *satisface* o *difiere en una serie en el enésimo lugar a lo sumo*. Y por otro lado, los términos teóricos sintácticos utilizados en las descripciones de expresiones o de construcciones cuyas propiedades semánticas están especificadas por las reglas.

Puede parecer enteramente claro que ningún hablante ordinario de una lengua natural tenga un dominio de los conceptos ordinarios. Pero, en realidad, esta consideración no muestra nada y en todo caso puede no ser pertinente. No muestra nada porque las reglas semánticas que consideran los

filósofos se expresan como axiomas de una teoría formal aplicada a versiones sometidas a un régimen o formalizadas de oraciones de la lengua en cuestión. De manera que saber qué conceptos nos incumben depende de la elección del régimen y de la teoría formal. En consecuencia, pretender que los hablantes ordinarios no tienen dominio de los conceptos teóricos semánticos puede ser el resultado de la elección del formalismo.

Además, estas consideraciones sobre la conceptualización pueden no ser pertinentes, pues si el conocimiento tácito de reglas semánticas se da por obra de la presencia de un procesador antes que por un estado representacional, luego la manera en que ese conocimiento se clasifique es un punto secundario. Todo dependerá de establecer si las descripciones de las expresiones que están al alcance de las reglas son conceptualizadas por la persona en cuestión.

En suma, tratándose de la clasificación de nuestro conocimiento de reglas semánticas, las consideraciones sobre la conceptualización no son decisivas. Pero el propio Evans trata rara vez este punto de conceptualización y lo considera diferente de la observación de que el conocimiento que tiene un hablante de las reglas semánticas está inferencialmente aislado y no al servicio de un proyecto del sujeto, separado del proyecto de hablar y de comprender una lengua.

Campbell (1982) replica a esto poniendo en tela de juicio las razones por las que se sostiene que el conocimiento de las reglas semánticas es subdoxástico. Expone dos argumentos. El primero está destinado a minimizar la significación de la tesis de que el conocimiento de las reglas semánticas tiene un empleo sólo muy limitado. Campbell sostiene que hay casos claros de conocimiento ordinario en los que aquello que se conoce tiene aplicaciones solamente muy limitadas. Así, el simple hecho de que el conocimiento de las reglas semánticas tiene una aplicación limitada no puede mostrar que el conocimiento tácito sea otra cosa, en definitiva, que el conocimiento en el sentido corriente del término.

Este punto es evidentemente correcto: con seguridad hay tipos de conocimiento que en realidad se utilizan muy poco. Pero la tesis de Evans sostenía que la información contenida en una regla semántica no está “*ni siquiera potencialmente* al servicio de otro proyecto” (pág. 339, la bastardilla es mía). Lo que se sabe de la intriga de *Bleak House* (el ejemplo es de Campbell, pág. 26) está por lo menos potencialmente al servicio de numerosos

proyectos; esos hechos conocidos podrían desplegarse inferencialmente si el sujeto tuviera una creencia condicional que relacionara uno de esos hechos refinados con un asunto urgente cualquiera. En consecuencia, la cuestión importante —que Campbell no plantea— es saber si el conocimiento de las reglas semánticas está también él *potencialmente* al servicio de los diferentes proyectos del sujeto.

Stich (1978, págs. 508-509) da un ejemplo de integración inferencial que por lo menos a primera vista parece amenazar el primer argumento de Campbell:

“Supongamos que en el caso de una regla hipotética *r* llegáramos a creer que si *r*, luego Chomsky está seriamente equivocado. Supongamos además que de hecho *r* sea una de las reglas almacenadas de nuestro proceso de tratamiento del lenguaje. Esta creencia, agregada a la creencia de que Chomsky está seriamente equivocado, no conducirá a la creencia de que Chomsky está seriamente equivocado”.

En este ejemplo un sujeto cree un hecho condicional que vincula una regla lingüística con un tema que interesa al sujeto. Si la regla lingüística conocida está potencialmente al servicio de los proyectos del sujeto, como lo están los hechos relativos a *Bleak House*, luego deberíamos esperar que el sujeto llegue a la conclusión de que Chomsky está seriamente equivocado. Pero Stich nos dice, también seriamente, que esto no puede ocurrir.

Sin embargo aquí debemos tener en cuenta una vez más el hecho de que el conocimiento de una regla semántica podría darse por la presencia de un procesador. La regla del *modus ponens* está realizada ella misma, supongámoslo, por un procesador que cumple computaciones inferenciales sobre estados representacionales. Así, para dar un ejemplo práctico de aislamiento inferencial respecto de la red de creencias, debemos especificar un conocimiento tácito dado como estado representacional. Nuestro ejemplo anterior, el de la percepción de la voz durante un tiempo dado, apuntaba a satisfacer esta condición. Pero el conocimiento tácito de la regla *r* no la satisface de manera clara. Esencialmente el mismo problema que surge con el ejemplo de Stich podría plantearse sobre la distinción señalada por Chomsky entre conocer la regla *R* y saber que la regla *R* es una regla válida (1986, pág. 17).

Hay por consiguiente una dificultad de principio en la utilización del concepto de aislamiento inferencial como criterio para clasificar el conocimiento tácito de una regla en el dominio de lo subdoxástico. Podemos intentar emplear la falta de conceptualización como criterio; pero ese criterio sólo se aplica a descripciones de expresiones y de construcciones en virtud de las cuales tales expresiones están sometidas a la regla. Y en el caso de las reglas semánticas ese criterio no es decisivo. Según el primer argumento de Campbell, el conocimiento de reglas semánticas es un caso límite.

El segundo argumento de Campbell está destinado a mostrar que el conocimiento de reglas semánticas debe clasificarse definitivamente dentro del ámbito de las actitudes. A causa del fenómeno constante de la dependencia contextual que hay en una lengua natural, la comprensión de un enunciado exige el empleo de reglas estrictamente semánticas y de conocimientos no lingüísticos ordinarios. Por consiguiente, las reglas semánticas están inferencialmente integradas en el interior de la red de las creencias (pág. 26).

“Evans sostiene que el conocimiento lingüístico no puede combinarse con el pensamiento general de un sujeto a fin de producir nuevas creencias. Y sin embargo, el fenómeno de la dependencia contextual significa que el conocimiento lingüístico es a menudo insuficiente para comprender el discurso. Comprender un acto de lenguaje es típicamente algo que procede a la vez de conocimientos lingüísticos y de conocimientos no lingüísticos. No es posible delimitar de antemano la información pertinente. El llamado aislamiento inferencial del conocimiento semántico es, pues, un mito.”

Este argumento parte del hecho bien conocido de la interpenetración de los factores semánticos y pragmáticos en la comprensión, y a primera vista parece decisivo. Sin embargo el argumento muestra demasiado, como puede comprobarse si se considera el ejemplo del reconocimiento visual del objeto.

Reconozco un automóvil, que tiene cierto color o cierta forma, estacionado en cierto lugar y lo reconozco como el automóvil de la abuela. Este reconocimiento depende en parte de toda clase de creencias sobre el mundo, entre las cuales están

las que tienen que ver con los automóviles. Este reconocimiento depende también —a menos que la teoría científica actual esté enteramente equivocada— del proceso del sistema visual. De manera que el reconocimiento del automóvil de la abuela es, por un lado, el producto de un proceso de tratamiento de la información visual y, por otro, el producto de conocimientos corrientes, como por ejemplo el de que los automóviles no se mueven por sí solos una vez que están estacionados (legalmente). Pero este hecho en sí mismo no da ni siquiera el comienzo de una prueba de que las representaciones y computaciones que entran en juego en la visión primitiva de los colores y de las formas sean parte de la red de creencias, pues por el contrario continúan siendo los ejemplos más claros que haya del ámbito subdoxástico.

El argumento de Campbell no es en definitiva concluyente. Tal como está expuesto invita a replicar que es sólo el producto final de las aplicaciones de las reglas semánticas —sus consecuencias para la significación de oraciones totales— lo que entra en interacción con la información de fondo.

Teniendo en cuenta los argumentos de Evans y de Campbell, de todo esto se desprende que nuestro conocimiento de las generalizaciones semánticas sistemáticas es un caso límite en lo que se refiere a la distinción entre el dominio de las actitudes y el dominio de los estados subdoxásticos.

¿En qué medida se puede emplear la modularidad para establecer una frontera en el interior de este territorio que es él mismo fronterizo? Evidentemente, aquí se trata de un problema empírico que depende mucho de los resultados de los trabajos que se realizan en psicología lingüística experimental. Pero esta perspectiva no se ve fortalecida por el hecho de que la modularidad está definida atendiendo a una clase extendida de rasgos característicos, pues diferentes rasgos característicos de los módulos sugieren diferentes maneras de trazar los límites de la modularidad. Se puede observar este conflicto en el caso de la visión.

Considérese en primer término el encapsulamiento informacional. A veces parece a primera vista que un proceso psicológico no apela a toda la riqueza de la información disponible en un sistema cognitivo. En esas circunstancias, siempre podemos defender la idea de que existe un proceso modular. Sostenemos que el punto de interacción no es el proceso desarro-

lado en el interior del módulo, sino que se trata antes bien de la información de salida del módulo. En otras palabras, trazamos el límite del sistema periférico modular antes en el proceso total. Cuanto menos se supone que el sistema periférico mismo trata la información, más fácil es sostener que ese sistema está informacionalmente encapsulado.

De manera que no es sorprendente que al discutir el encapsulamiento informacional Fodor diga lo siguiente sobre el sistema visual (1983, pág. 74):

“Muchos datos confirman que hay efectos contextuales en ciertos aspectos del reconocimiento visual. Pero esos datos no tienen ninguna importancia en la presente discusión a menos que haya una razón independiente para creer que esos aspectos del reconocimiento de objetos forman parte del análisis visual de entrada. Quizás el sistema periférico en el caso de la visión especifique el estímulo tan sólo como ‘esbozos primarios’...”

Considerando que el esbozo primario es un estado que se sitúa al comienzo mismo del tratamiento de la información en el modelo de Marr, esta sugestión sobre los límites del sistema periférico no puede aplicarse a una gran parte de lo que razonablemente se puede llamar el tratamiento de información del sistema central.

Por otro lado, mientras se supone que los niveles intermedios de un módulo son relativamente inaccesibles al sistema central, se supone que las informaciones de salida de un módulo son “fenomenológicamente salientes”(pág. 87).

“Me parece que tenemos necesidad de un concepto de proceso perceptual que haga que los productos de la percepción puedan figurar como premisas de decisiones conscientes y de inferencias”.

Esto concuerda muy bien con la imagen de los sistemas periféricos que presentan al sujeto pensante candidatos para la creencia. Teniendo en cuenta esta imagen, podemos considerar los efectos de contextos como algo que constituye solamente un aspecto del proceso intelectual de producción de creencias a la

luz de conocimientos de fondo o de segundo plano, antes de que se formen juicios.

Respondiendo a esta idea Fodor afirma (pág. 94):

“Diversos candidatos [para las salidas del procesador visual] deben ser rechazados a causa de la inaccesibilidad fenomenológica. Pienso en representaciones tales como los esbozos ‘primarios’, ‘2,5D’ y ‘3D’ de Marr... Semejantes representaciones parecen demasiado endebles. Si las aceptamos como capaces de definir las salidas del procesador visual, deberemos decir que todo reconocimiento de objeto no es, en rigor de verdad, un fenómeno de la percepción visual pues, en esos niveles de representación, sólo son especificadas ciertas propiedades geométricas. Pero, desde el punto de vista de la accesibilidad fenomenológica, es seguro que la percepción es ante todo un reconocimiento de objetos y de hechos”.

Aquí, con la preocupación de defender el carácter saliente fenomenológico de las salidas de un módulo, comprobamos que Fodor atribuye gran parte del tratamiento de la información al sistema periférico.

El resultado neto de estas presiones conflictivas ejercidas en las fronteras de la modularidad es el de que los esbozos 2,5D y 3D de Marr entran dentro de los casos límites. Esto no refuerza la perspectiva de utilizar la modularidad para fijar la clasificación de los casos límites delicados, ya en la esfera de lo subdóxico, ya en la de las actitudes.

En el caso del tratamiento de la información lingüística, se da una indeterminación similar debida a los diferentes rasgos característicos de la modularidad. Por un lado, la condición de encapsulamiento informacional sugiere que una gran parte de lo que razonablemente puede llamarse tratamiento de información no es en realidad realizado en modo alguno por el sistema lingüístico. Descubrir la ironía no es evidentemente un proceso encapsulado (pág. 88); nadie pensará en ponerlo en tela de juicio. Pero Fodor sugiere de manera más radical que la tarea del sistema lingüístico podría ser sencillamente la de clasificar un carácter (*token*) que se presenta como parte de un tipo sintáctico y tal vez asignarle una forma lógica (pág. 89).

“En conclusión, si buscamos una propiedad interesante de las enunciaciones que pueda realizarse por obra de los sistemas rígidamente encapsulados —es decir, una propiedad que hasta pueda ser calculada esencialmente de arriba a abajo (*bottom to top*) por procesadores—, entonces la identidad genérica de la enunciación junto tal vez con su forma lógica parece ser el candidato que se impone.”

Esto deja sólo un papel comparativamente reducido al sistema lingüístico. Pero, por otro lado, se supone que la información de salida del sistema lingüístico es fenomenológicamente saliente. Y como lo observa Fodor, cuando uno oye enunciar a alguien algo en una lengua que comprende, no puede literalmente evitar oír lo que dice el hablante: “Y es lo que se *dice* lo que no se puede evitar oír, no lo que es simplemente *enunciado*” (pág. 55). Si la información de salida del sistema periférico es fenomenológicamente saliente, luego no es sólo la sintaxis o la forma lógica lo que calcula el sistema lingüístico sino que es también el contenido o el mensaje.

En el caso del tratamiento de la información visual, las presiones conflictivas que pesan sobre las fronteras de la modularidad dejan indeterminada la índole de cada componente del reconocimiento visual de los objetos después del esbozo primario. También en el caso del tratamiento de la información lingüística, esas mismas presiones conflictivas dejan sin determinar la índole de cada componente de la comprensión de los enunciados después de la asignación de la forma lógica. En particular, lo que queda indeterminado por todo esto es la índole del tratamiento sistemático de la información semántica.

Al sugerir que la modularidad es la esencia real de la subdoxasticidad, esperábamos principalmente fijar la condición del conocimiento de las reglas semánticas, pues se trataba de un caso límite delicado una vez aplicados los criterios filosóficos. En la medida en que la modularidad se caracteriza por una clase extensa de rasgos, esa esperanza parece vana. Permanentemente tenemos necesidad de un análisis de la modularidad que nos revele por qué sus señales características acompañan el encapsulamiento informacional, que es la esencia de la modularidad.

Hemos comenzado con la cuestión de establecer si los estados de conocimiento tácito eran semejantes a los estados de actitudes proposicionales o diferentes de ellos. En la sección 1, hemos examinado algunas de las complejidades de la respuesta de Chomsky a la cuestión. En términos generales, su tesis es la de que una vez que tiene uno en cuenta la distinción entre el hecho de conocer la regla R y el hecho de saber que la regla R es válida, no hay diferencia importante entre los estados de actitudes y los estados de conocimiento tácito. En la sección 2, presentamos tres marcas intuitivas de distinción entre las creencias y ejemplos de estados subdoxásticos, incluso el conocimiento tácito de reglas sintácticas y estados que intervienen en las fases iniciales del tratamiento de la información perceptual. En la sección 3, hemos expuesto tres razones para decir por qué, en ese contexto, una noción indiferenciada de accesibilidad a la conciencia no permite explicar nada; así hemos justificado uno de los elementos del escepticismo de Chomsky en cuanto a la importancia de la distinción entre conocimiento consciente (o creencia) y conocimiento tácito. De paso introdujimos un factor de complicación importante, a saber, la distinción entre estados representacionales y procesos computacionales.

Una vez descartada la accesibilidad a la conciencia, quedan dos marcas intuitivas. Una era la diferencia entre la integración inferencial y el aislamiento inferencial. La otra era la diferencia entre estados que están conceptualizados por el sujeto y estados que no lo están. En la sección 4, hemos seguido la línea de indagación sugerida por la distinción entre estados conceptualizados y estados no conceptualizados. Esto nos ha llevado a la idea de que una distinción fundada entre las creencias y los estados subdoxásticos debe apoyarse en una distinción entre dos clases de contenido: uno conceptualizado y sujeto a la condición de generalidad y el otro no conceptualizado y sin estar sujeto a esa condición. La cuestión de clasificar cierta clase de conocimiento tácito se hace así una cuestión derivada, pues los procesos se clasificarían entonces en función de sus entradas y sus salidas. En la sección 5, hemos seguido otra vía de indagación sugerida por uno de los componentes de la idea de aislamiento inferencial. Lo que se ha desprendido de esta sección era la sugestión de que la modularidad de Fodor puede

ser la esencia real del concepto de subdoxasticidad del filósofo. El valor de esta proposición depende de la capacidad explicativa de la concepción de psicología científica de Fodor. Pero depende también de la posibilidad de suministrar explicaciones (desde el punto de vista de la modularidad) de las marcas intuitivas de la subdoxasticidad. Y depende asimismo del éxito que obtengamos al utilizar la noción científica para explicar casos límites de la subdoxasticidad.

En la sección 6, hicimos notar hasta qué punto las consideraciones relativas a la conceptualización y al aislamiento inferencial dejan indeterminada la índole del conocimiento tácito de reglas semánticas y a fortiori cuando se tiene en cuenta la distinción entre estados representacionales y procesos computacionales. Este es el tipo de casos en que sería útil remitirse a la ciencia para decidir sobre la índole del conocimiento de las reglas semánticas en función de la posición de los límites de la modularidad. Sin embargo, el hecho de que la modularidad se caracteriza por una clase muy amplia de rasgos constituye aquí una dificultad grande. Y esos diferentes límites posibles dejan los casos límites tan indeterminados como estaban antes.

### Referencias bibliográficas

- Ayers, M., 1981, "Locke versus Aristotle on natural kinds", *Journal of philosophy*, 78.
- Baker, G. P., y Hacker, P.M.S., 1984, *Language, sense and nonsense*, Blackwell.
- Campbell, J., "Knowledge and understanding", *Philosophical Quarterly* 32.
- Chomsky, N., 1965, *Aspects of the theory of syntax*, MIT Press, tr. fr. Seuil 1967 [Hay versión castellana: *Aspectos de la teoría de la sintaxis* Madrid, Aguilar, 1976.]
- 1976, *Reflexions on language*, Fontana, Collins, tr. fr. Maspero, 1978. [Hay versión castellana: *Reflexiones sobre el lenguaje* Barcelona, Ariel, 1979.]
- 1980, *Rules and representantions*, Blackwell, tr. fr. Maspero, 1978.
- 1986, *Knowledge of language: its nature, origin, and use*. Praeger.
- Davies, M., 1986, « Tacit knowledge, and the structure of thought and language », en Travis (comp.), *Meaning and interpretation*, Blackwell.
- 1987, "Tacit knowledge and semantic theory: Can a five per cent difference matter?" *Mind*, 96.
- Evans, G., 1981, "Semantic theory and tacit knowledge", en *Collected papers* (Oxford University Press, 1986).
- 1982, *The varieties of reference*, Oxford University Press.

Fodor, J., 1983, *The modularity of mind*, MIT Press, tr. fr. Minuit, 1985.

Quine, W.V.O., 1972, "Methodological reflections on current linguistic theory", en Davidson y Harman (comps.), *Semantics of natural lan-guage*, Reidel.

Stich, S., 1978, "Beliefs and subdoxastic states", *Philosophy of science*, 45.



## ¿Tienen efectos físicos los fenómenos mentales? (¿ Es el monismo anomal una forma de epifenomenismo?)

Daniel Laurier

Este artículo está organizado del modo siguiente. Comienzo diciendo lo que es el monismo anomal y proponiendo una reconstrucción de la argumentación desarrollada por Davidson (1980) con miras a establecerlo. Luego hago algunos comentarios destinados a precisar el alcance de las dos primeras premisas de la argumentación; por fin examino algunas objeciones tendientes a mostrar que el monismo anomal es incapaz de justificar la eficacia causal de los fenómenos mentales.

### 1. Reconstrucción del argumento

El monismo anomal es la conjunción de las tesis siguientes:

TI (*token-identidad*): todo evento mental es idéntico a un suceso físico (1980: 231, 248).

AM (anomalismo de lo mental): no puede haber leyes estrictas que permitan explicar o predecir eventos mentales (1980: 208).

La argumentación general de Davidson a fin de establecer estas dos tesis se desarrolla en tres grandes etapas. Davidson utiliza primero la tesis AM como premisa para establecer tanto la propia TI (1980: 231, 248) como una versión restringida de TI (1980: 224). Trata entonces de justificar AM apoyándose (entre

otras premisas) en una tesis más restringida, la de que no puede haber leyes *psicofísicas* estrictas. Y en una tercera etapa debe establecer este último punto; para hacerlo invoca el hecho de que la aplicación de los conceptos psicológicos presupone normas de racionalidad que serían ajenas a la esfera de la física.

En este artículo sólo me intereso por la primera etapa de este razonamiento, es decir, la que tiende a demostrar la tesis de la *token-identidad* al suponer el anomalismo de lo mental. De manera que no hemos de tratar aquí esta última tesis.

El argumento más explícito expuesto por Davidson aparece en "Mental Events" (1980: 224), pero sólo tiene como conclusión una versión restringida de la tesis de la identidad, a saber, la tesis TIR, según la cual todo evento mental que esté causalmente relacionado con un suceso físico es un hecho físico. El argumento parece ser el siguiente:

1) ICR (principio de interacción causal restringida): por lo menos ciertos eventos mentales son causas o efectos de sucesos físicos (1980: 208).

2) CNC (principio del carácter nomológico de la causalidad): dos eventos que tienen una relación causal están sujetos a una ley estricta (1980: 208).

3) AM.

4) Todo hecho que está sometido a una ley física estricta tiene una descripción física.

5) Todo hecho que tiene una descripción física es un hecho físico.

TIR (*token-identidad restringida*).

Se ve que si TIR es una consecuencia de 1)-5), entonces TI será una consecuencia del argumento obtenido al reemplazar la premisa 1) por la siguiente:

6) IC (principio de interacción causal): todo evento mental es una causa o efecto de un suceso físico.

Davidson parece adherirse precisamente a IC y deducir TI de 2)-6) en "Philosophy as Psychology" (1980: 231) y en "The Material Mind" (1980: 248).

Davidson razona de la manera siguiente. Sea *m* un evento mental que tiene una relación causal con un hecho físico. En virtud del carácter nomológico de la causalidad (es decir, la premisa 2)), *m* está sujeto a una ley estricta. Pero a causa del

anomalismo de lo mental —es decir a causa de la premisa 3)— esa ley sólo puede ser una ley física estricta y en virtud de la premisa 4) se sigue que *m* tiene una descripción física y en virtud de la premisa 5) que *m* es pues un suceso físico. Por consiguiente, todo evento mental que esté relacionado causalmente con un suceso físico es él mismo un suceso físico (=TIR).

Este razonamiento no es válido a causa de tres razones por lo menos.

En primer lugar, parece que el razonamiento no puede ser exactamente el mismo si el evento *m* es una causa o si es un efecto de un suceso físico, pues todo hecho sujeto a una ley estricta sólo puede explicarse o predecirse por medio de una ley estricta, ya que todo suceso que puede explicarse o predecirse mediante una ley estricta está sujeto a esa ley estricta. Un evento sólo se explica o se predice mediante una ley si ese evento ejemplifica el consecuente. Ahora bien, es perfectamente concebible que un evento mental cause sucesos físicos sin ser el efecto de un suceso físico o que sólo ejemplifique el antecedente de una ley estricta.

Esto me lleva a considerar la segunda dificultad planteada por el razonamiento de Davidson (enseguida volveré a considerar la primera). Supongamos que el evento mental *m* esté causado por un hecho físico. Deberíamos poder llegar a la conclusión (en virtud del carácter nomológico de la causalidad) de que ese evento está sujeto a una ley estricta y por consiguiente que debe poder explicarse o predecirse mediante una ley estricta, lo cual está en contradicción con el anomalismo de lo mental. Como la intención de Davidson es admitir que un evento mental puede estar sujeto a una ley física estricta, es evidente que la tesis del anomalismo de lo mental tiene que reformularse para que llegue a ser compatible con esta conclusión.

Hay que recordar que Davidson considera las leyes como enunciados, es decir, expresiones lingüísticas; esto quiere decir que un evento está sujeto a cierta ley únicamente si satisface un predicado (una descripción que contenga un predicado) que aparece en esa ley en una descripción dada, y no la noción absoluta de ejemplificación de una ley *simpliciter*.

Si un evento satisface un predicado que aparece en una ley estricta que permite explicarlo o predecirlo, se puede decir entonces que esa ley permite explicar o predecir dicho evento *en*

*la medida en que satisface ese predicado*. Precisos estos puntos parece que la tesis del anomalismo mental (en la formulación de Davidson) equivale a afirmar que no puede haber descripción en la que una ley estricta permita explicar o predecir un evento mental, lo cual es manifiestamente incompatible con la intención de Davidson. Por consiguiente, sugiero reformular esa tesis de la manera siguiente:

AM': no puede haber leyes estrictas que permitan explicar o predecir un evento mental *en cuanto mental* (o de manera equivalente, no puede haber descripción mental en la que una ley estricta permita explicar o predecir un evento mental).

Esta reformulación del anomalismo de lo mental permite eliminar la segunda dificultad planteada por el razonamiento de Davidson. Supongamos, en efecto, que el evento mental *m* esté *causado* por un hecho físico; entonces podemos razonar de la manera siguiente. El evento *m* no puede explicarse ni predecirse por medio de una ley estricta, pero en virtud del anomalismo de lo mental (del modo en que acabamos de reformularlo), esa ley estricta debe ser una ley física estricta y por lo tanto *m* debe ser un suceso físico.

Sin embargo (y aquí torno a considerar la primera dificultad), si *m* *causa* un suceso físico (sin ser causado por un suceso físico), luego no se puede concluir que *m* pueda explicarse o predecirse por medio de una ley estricta y por consiguiente no se puede llevar a su término el razonamiento. Para hacerlo se necesita una formulación más general de la tesis del anomalismo, y yo propongo ésta:

AM'': no puede haber leyes estrictas que rijan un evento mental *en cuanto mental* (es decir: un evento no puede ejemplificar una ley estricta en una descripción mental).

El razonamiento se desarrolla entonces del modo siguiente. Sea *m* un evento mental causalmente relacionado con un suceso físico. En virtud del carácter nomológico de la causalidad, *m* está sujeto a una ley estricta. En virtud del anomalismo de lo mental (según la nueva formulación), esa ley no puede ser sino

una ley física estricta. Síguese de ello (por la premisa 4) que *m* tiene una descripción física y (por la premisa 5)) que es un hecho físico.

Este razonamiento no se ve afectado por el hecho de que *m* sea una causa o un efecto de un suceso físico, pero el razonamiento no es todavía válido puesto que (y ésta es la tercera dificultad planteada por el razonamiento original) el anomalismo de lo mental, ni siquiera en la nueva formulación, permite concluir que

7) si un evento mental está sujeto a una ley estricta, ésta debe ser una ley *física* estricta.

Aparentemente se obtendría un razonamiento válido agregando 7) como premisa suplementaria, pero AM" sería entonces superfluo puesto que parece ser una consecuencia de 7). Si en efecto se conviene en que

8) una descripción física no es una descripción mental, lo que en modo alguno es problemático, luego 7) implica que si un evento mental está sujeto a una ley estricta no puede ejemplificar esa ley en una descripción mental.

Pero se obtiene también un razonamiento válido y aparentemente más cercano al de Davidson conservando AM" y agregando como premisa la conversión de la premisa 8), a saber:

9) toda descripción no mental es una descripción física.

Propongo, pues, reconstruir la argumentación de Davidson del modo siguiente 1), 2), AM", 9), 4), 5), TIR.

Las premisas importantes para mis fines son las tres premisas explícitas de Davidson, esto es, el anomalismo de lo mental, el carácter nomológico de la causalidad y el principio de interacción causal. Ya he dicho algo sobre la interpretación de la tesis del anomalismo de lo mental. De manera que ahora basta con comentar brevemente las otras dos premisas para poder luego ocuparme de las objeciones que me interesan.

Para comprender bien el principio del carácter nomológico de la causalidad (= 2) = CNC) hay que recordar que Davidson considera la causalidad como una relación entre sucesos con-

cebidos como sucesos particulares (es decir, como no repetibles). Como Davidson admite una ontología de hechos singulares puede tratar el contexto "x causa y" como un contexto plenamente extensional. El principio CNC significa pues que, si un evento *a* causa un evento *b* (o es causado por éste), entonces hay una descripción de *a* y una descripción de *b* en las cuales *a* y *b* ejemplifican una ley estricta o, en otras palabras, hay una ley estricta que relaciona *tipos* de eventos a los cuales pertenecen respectivamente *a* y *b*. Este principio no implica que si *a* y *b* son descripciones de eventos causalmente relacionados haya una ley estricta que contiene *a* y *b*, es decir, una ley estricta que abarque *a* y *b*, si éstos satisfacen respectivamente las descripciones *a* y *b*. Davidson, que yo sepa, no da ningún argumento positivo en favor de CNC, sino que simplemente lo propone (EAE: 160) como una manera de conciliar dos concepciones aparentemente incompatibles de causalidad, a saber, la concepción de que una relación causal implica una relación nomológica (Hume, Mill) y la concepción según la cual un enunciado causal similar (es decir, un enunciado de la forma "x causa y") no implica ninguna enunciación nomológica (Ducasse). La conciliación se debe al hecho de que en virtud de CNC una relación causal implica la existencia de una ley sin implicar ninguna ley particular.

El principio de interacción causal restringida (=1) = ICR) es a priori muy plausible y, por lo demás, está ampliamente aceptado. Pero esto quizá sólo se deba a la convicción que tenemos de que por lo menos ciertas sensaciones causan hechos físicos y son causadas por ellos. El hecho de extender la mano sobre una llama causará una sensación de dolor que a su vez (normalmente) provocará un grito o una contracción muscular o ambas. De modo que el principio ICR no es en modo alguno problemático si se considera toda la gama de los eventos mentales posibles, pero esto ya no ocurre cuando limitamos la clase de los eventos mentales pertinentes a los eventos intencionales. Ahora bien, Davidson tan pronto limita su argumentación a los eventos mentales intencionales (1980: 229-230; 246), tan pronto considera todos los eventos mentales como eventos intencionales (1980: 210-212). En el segundo caso, el principio ICR no parece necesitar ninguna argumentación particular, pero asimilar todos los eventos mentales a fenómenos intencionales es suficientemente problemático y artificial

para que Davidson se sienta obligado a establecer que ciertos fenómenos (que son indiscutiblemente intencionales) tienen relaciones causales con fenómenos físicos. Esto importa tanto más cuanto que el argumento destinado a establecer el anomalismo de lo mental sólo es aplicable a los fenómenos intencionales. Es principalmente en "Actions, Reasons and Causes" donde Davidson intenta resolver la cuestión.

Sostiene allí que el concepto de causalidad debe invocarse para mostrar la diferencia entre "tener una razón para obrar" y "obrar por tal o cual razón". Davidson trata en este artículo de dilucidar la naturaleza y el funcionamiento de este tipo familiar de explicar la acción, que consiste en citar la razón (o una razón) por la cual se realizó una acción dada y que Davidson llama una racionalización. Sin examinar aquí todo el análisis de Davidson, importa precisar sus caracteres fundamentales.

Davidson distingue primero, dentro de la clase de las racionalizaciones, aquella que explica una acción mencionando una razón primaria por la cual se realizó esa acción. Llamaré a esas racionalizaciones racionalizaciones primarias. De suerte que una racionalización primaria es una racionalización que invoca una razón primaria y una razón primaria es un par de estados intencionales, uno de los cuales es una proactitud (es decir, una actitud tal como desear, querer, anhelar, etc. respecto de toda acción de cierto tipo) y el otro es una creencia de que cierta acción es de ese tipo (o también una creencia de que cierta acción es de un tipo ante el cual el agente tiene una proactitud).

No toda racionalización es una racionalización primaria puesto que se puede racionalizar una acción citando, por ejemplo, sólo un componente o dos componentes de una razón primaria o más generalmente invocando los fines, las disposiciones, los sentimientos o las intenciones del agente. Davidson sostiene sin embargo que toda acción tiene por lo menos una racionalización primaria y que para comprender cómo una razón cualquiera puede racionalizar una determinada acción (es decir, contar como una razón *por la cual* se realizó esa acción) es necesario y suficiente ver, en grandes líneas, cómo construir una racionalización primaria de esa acción partiendo de esa razón. El problema general que plantea Davidson, es decir, el problema de saber en qué condiciones una razón racionaliza una acción dada, se reduce pues al problema de establecer en

qué condiciones una razón primaria racionaliza una acción dada.

Davidson observa primero que una racionalización (y a fortiori una racionalización primaria) sólo explica una acción en una descripción, así como una ley sólo rige (y por lo tanto sólo explica) un hecho en una descripción. Por lo demás, los dos casos no dejan de tener cierta relación, puesto que Davidson considera las acciones como eventos de cierta clase y las racionalizaciones como explicaciones de cierta clase. Ahora bien, según él, las explicaciones (causales, nomológico-deductivas o racionalizantes) son construcciones lingüísticas no extensionales y por consiguiente sensibles a la manera en que son descritos los fenómenos en cuestión. Por ejemplo, la amplitud de la dilatación de una columna de mercurio puede explicarse por el aumento de temperatura y una ley que relacione la temperatura y el volumen de una masa de mercurio. La dilatación de una columna de mercurio en un momento dado puede ser idéntica a la dilatación de cierta masa de líquido rojo pero no se puede explicar la amplitud de la dilatación de una masa de líquido rojo invocando la ley que relaciona la temperatura y el volumen de una masa de mercurio; y hasta puede no haber ninguna ley pertinente. Del mismo modo, es posible explicar por qué un agente movió un conmutador diciendo que deseaba iluminar la habitación. El ademán de mover el conmutador puede ser también el mismo que pone en marcha un sistema de alarma. Pero manifiestamente no se puede explicar por qué el agente puso en marcha el sistema de alarma diciendo que deseaba iluminar la habitación. Ese deseo sólo explica la acción dada como acción de mover el conmutador y no como acción de dar la alarma.

Teniendo en cuenta estas aclaraciones, para Davidson se trata entonces de decir en qué condiciones una razón primaria racionaliza una acción determinada en una descripción dada. Davidson propone dos condiciones necesarias pero no suficientes. Una razón primaria  $R = (x, y)$  es una razón por la que un agente realiza la acción  $A$  en la descripción  $d$  solamente si

10)  $x$  es una proactitud del agente con respecto a acciones de cierto tipo y si  $y$  es una creencia del agente de que  $A$ , en la descripción  $d$ , es de ese tipo.

11)  $R$  es la causa de  $A$ .

La condición 11) se hace necesaria por el hecho de que un agente puede muy bien tener una razón (y una razón primaria) para realizar cierta acción y sin embargo no realizarla o no realizarla *por* esa razón. En este caso, su acción no queda explicada por dicha razón; una acción sólo es racionalizada por una razón que es su causa. Davidson sostiene que las racionalizaciones son un tipo de explicación causal. Poco después me ocuparé de la razón por la cual 10) y 11) no son suficientes para que R racionalice A. Por el momento es más oportuno preguntarse en qué 11) contribuye a establecer el principio ICR.

En efecto, si no se puede dudar de que las creencias y las proactitudes son fenómenos mentales, no resulta claro que sean eventos; y si es claro que las acciones son eventos puede dudarse de que sean fenómenos físicos. Pero si las creencias o las proactitudes no son eventos, no pueden (según la concepción davidsoniana de la causalidad) tener relación causal con nada. Davidson concede sin embargo (1980: 12) que en rigor de verdad no es una creencia (una proactitud o una razón) lo que causa una acción sino que es su aparición (activación, adquisición o realización). Ahora bien, la misma creencia (proactitud, razón) puede aparecer (ser activada o realizada) más de una vez; hay pues que comprender que es *la* aparición de una razón primaria  $R = (x,y)$  lo que causa la acción y no esa razón misma como tal.

Faltaría precisar cuál es exactamente la relación que hay entre una razón primaria y las actitudes que la componen o, más exactamente, entre la aparición de una razón primaria y las apariciones de las actitudes que la componen. Davidson deja esta relación en la oscuridad, pero me parece que una aparición de  $x$  o de  $y$  causa la acción A, pues de otra manera no se podría decir que la aparición de la razón primaria  $R = (x,y)$  causa A. El examen de esta cuestión puede arrastrarnos a una discusión de los criterios de individuación de los eventos que será mejor diferir. Yo propongo simplemente suponer de manera provisional que una aparición de  $R = (x,y)$  causa A sólo si una aparición de  $x$  causa A o una aparición de  $y$  causa A. Para poder llegar a la conclusión de que un evento mental puede causar un evento físico partiendo del hecho de que la aparición de una razón primaria puede causar una acción, hay que suponer por lo menos que ciertas acciones causadas por la aparición de una razón primaria son idénticas a hechos físicos o causan hechos

físicos. Si las acciones deben entenderse como eventos mentales, sería evidentemente un razonamiento circular en el presente contexto presuponer que las acciones en cuestión son también sucesos físicos, puesto que se trata precisamente de establecer una premisa que permita concluir que ciertos eventos mentales son hechos físicos. Por otro lado, no se ve bien por qué, en esta hipótesis, es necesario mostrar que la aparición de una razón primaria puede causar una acción para establecer el principio ICR, ya que sería necesario y suficiente mostrar (lo cual parece evidente) que ciertas acciones no son eventos mentales, y el principio ICR es una consecuencia inmediata del hecho de que la aparición de una razón primaria puede causar una acción. Este podría ser un motivo para pensar que Davidson no considera las acciones como eventos mentales; y podríamos encontrar otro motivo en el hecho de que los predicados o descripciones de acciones no satisfacen normalmente el criterio de lo psicológico propuesto por Davidson en "Mental Events". Sin embargo, resulta molesto comprobar que en varios pasajes (1980: 230, 240, 246, 249) Davidson incluye explícitamente los predicados de acción en la clase de los predicados psicológicos. Por otra parte, parece que hay toda una categoría de acciones que podríamos llamar "no corporales" y de las cuales difícilmente podría negarse que sean eventos mentales, como por ejemplo la acción de calcular mentalmente, de deliberar, de tomar una decisión, de imaginar, etc... Si se trata aquí de acciones (y admito que podría quererse limitar la clase de las acciones a la clase que yo llamo acciones corporales) no se ve bien cómo evitar la conclusión de que por lo menos ciertas acciones son eventos mentales; ahora bien, podría parecer extraño que ciertas acciones sean eventos mentales y otras no lo sean.

Con todo eso por el momento parece inútil proseguir esta discusión ya que lo principal, es decir, ICR, parece adquirido cualquiera que sea la índole última de las acciones, si uno admite que la aparición de una razón puede causar una acción corporal. Supongamos que, en efecto, una aparición de R cause la acción A. Si A no es un evento mental, es un hecho físico (no confundir esto con 9)), y por consiguiente la aparición de R causa un suceso físico. Si A es una acción corporal y un evento mental, A debe tener efectos físicos. Esto permite llegar a la conclusión de que ciertos eventos mentales causan sucesos físicos y por lo tanto que ciertos eventos mentales son causas o efectos de

hechos físicos. Hemos de observar que Davidson no da ningún argumento en favor de la circunstancia de que ciertos eventos mentales están causados por hechos físicos, aunque admite explícitamente (1980: 208) una concepción causal de la percepción y por tanto una forma un poco más completa del principio ICR, a saber, el principio de que ciertos eventos mentales son causas de hechos físicos y que ciertos eventos mentales son efectos de hechos físicos.

Si las demás premisas que hemos identificado son aceptables, el argumento de Davidson permite pues concluir que todos los eventos mentales que causan un hecho físico y que todos los eventos mentales que son causados por un hecho físico son idénticos a sucesos físicos.

## 2. Respuesta a algunas objeciones

He presentado todos los elementos necesarios para discutir útilmente una objeción planteada por varios autores (Stoutland 1980, 1982, 1985, Honderich 1982, 1983, 1984, Follesdal 1985, Kim 1984, Robinson 1982) contra la posición de Davidson. Todos estos autores no formulan sus críticas en los mismos términos, pero en lo esencial parecen estar de acuerdo en reprochar a Davidson el no dar cuenta de la eficacia causal de lo mental como tal <sup>2</sup> y de hacer derivar esta dificultad de la conjunción de los tres principios IC, CNC y AM. Tendré principalmente en cuenta en mi exposición las versiones respectivas de Honderich y de Stoutland.

Primero me ocuparé de Honderich, que es el más explícito y aparentemente el más radical, pues no vacila en blandir el espectro del epifenomenismo. Muestra primero que no son todas las características de dos hechos en relación causal las características causalmente pertinentes. Así, cuando se coloca un kilo de manzanas rojas en una balanza de fabricación francesa no es el color rojo, sino que es el peso de las manzanas lo pertinente para explicar el hecho de que la balanza marque un kilo, y no es el hecho de que una balanza francesa marque un kilo lo que queda explicado. Esto es perfectamente compatible con la tesis de Davidson según la cual un hecho conserva las mismas relaciones causales cualquiera que sea la manera en que ese hecho esté descrito. Honderich no pretende que un hecho no cause otro si satisface cierta descripción sino que

simplemente subraya que, entre todas las descripciones de un suceso, sólo algunas son pertinentes para explicar algunas de sus relaciones causales. De manera que se puede afirmar (sin oponerse a Davidson) de un evento  $e$  en una descripción  $d$  que es una causa (o un efecto) de un evento  $e'$  en cierta descripción  $d'$ , si se entiende que es una causa o un efecto de  $e'$  y que  $d$  y  $d'$  son causalmente pertinentes (respecto de esta relación causal). Ahora bien, parece que ése es el sentido que Davidson da a la expresión “ $d$  explica (o es explicado) causalmente (por)  $d'$ ”. Davidson sostiene, en efecto, que la relación de explicación causal, a diferencia de la relación de causalidad, es intencional (1980: 154).

Una vez admitido el concepto de rasgo causalmente pertinente, Honderich presenta entonces el principio del carácter homológico de los rasgos causalmente pertinentes (CNTCP)<sup>3</sup>, que toda teoría adecuada de la causalidad debe satisfacer, según él, y que se puede formular así:

13) Si  $e$  en una descripción  $d$  es una causa o un efecto de  $e'$  en una descripción  $d'$ , entonces  $e$  y  $e'$  ejemplifican una ley en las descripciones  $d$  y  $d'$  respectivamente.

Esto equivale a decir en términos más naturales que todo rasgo causalmente pertinente de un evento entra dentro de una ley (es decir, es un rasgo “nómico”).

El monismo anomal se encuentra pues en un dilema. En efecto, si se admite ahora que un suceso físico como mental (es decir, si se admite que las características mentales de un evento mental pueden ser causalmente pertinentes), luego CNTCP e IC implican juntos la existencia de leyes psicofísicas y por consiguiente la falsedad de AM” y del monismo anomal. Si por el contrario uno niega que un evento mental pueda tener una relación causal con un evento físico en virtud de sus características mentales, parece que el monismo anomal (en conjunción con CNTCP) tiene como consecuencia el hecho de que los fenómenos mentales no tienen como tales ninguna eficacia causal. El monismo anomal, si es consecuente, se basaría así en una interpretación tendenciosa y perversa del principio de interacción causal.

Si bien Honderich sugiere que el monismo anomal no es en definitiva más que una forma de epifenomenismo, resulta

evidente que no se propone negar que el monismo anomal sea compatible con la existencia de interacciones causales entre eventos mentales y eventos físicos, sino que solamente niega que sea compatible con la *pertinencia causal* de las características mentales. Por consiguiente parece que el término “epifenomenismo” está mal elegido, puesto que en el sentido habitual el epifenomenismo pretende que un evento mental no puede causar un suceso físico. Y tanto más cuanto que la pertinencia causal se les niega no sólo a las características mentales de las causas mentales sino también a las características de los efectos mentales, de suerte que el mismo tipo de abuso de lenguaje que lleva a Honderich a hablar de “epifenomenismo” debería más bien llevarlo a hablar de “paralelismo”.

A primera vista sólo hay dos maneras de responder a esta objeción sin renunciar al monismo anomal (o a las premisas de que éste depende). La primera, que parece contar con el favor de los discípulos de Davidson, consiste simplemente en decir que esa objeción no es una objeción, puesto que no son los tipos de eventos o sus características los que tienen una eficacia causal, sino que son los eventos en cuanto *tokens*. Si un evento mental *m* que es idéntico a un evento físico *p*, causa otro evento físico *q*, no se puede pretender que esto sea un “simple accidente”, pues todo el mundo admite que *m* puede estar nomológicamente relacionado con *q* y *p* en la descripción *m* (o en cualquier otra descripción mental). Pero si es este el sentido de la objeción de Honderich (y éste se expresa precisamente de esa manera en 1983: 148), es claro, como lo observa Smith (1984: 85), que dicha objeción equivale sencillamente a negar el anomalismo de lo mental o a pretender que éste está en conflicto con el sentido común o a ambas cosas. Pero entonces parecería que el monista anomal pudiera permanecer perfectamente en su posición y negar que las características mentales tengan una pertinencia causal, en el sentido de Honderich, sin creerse por eso en oposición radical al sentido común, puesto que dista mucho de ser evidente por sí mismo que esto implique algo bien definido referente a la existencia de relaciones nomológicas entre lo mental y lo físico.

Sin ser absolutamente indefendible, esta actitud un poco dogmática no es sin embargo ni perfectamente satisfactoria ni la única compatible con el punto de vista de Davidson. En efecto,

éste podría admitir la pertinencia causal de las características mentales con la condición de rechazar el principio CNTCP. Semejante actitud parece, por otra parte, recomendada por la circunstancia de que Davidson sostiene a la vez que la explicación causal es una relación no extensional (1980: 154; 1985 b: 225-226) y que las racionalizaciones son explicaciones causales, pues estas dos tesis implican que las características mentales son causalmente pertinentes, en tanto que el anomalismo de lo mental excluye el principio CNTCP.<sup>4</sup> De manera que Davidson aparentemente debe negar que todos los rasgos causalmente pertinentes sean nómicos o, en otras palabras, que todas las explicaciones causales legítimas sean nomológicas.

Davidson es muy explícito sobre el hecho (ya señalado) de que una razón primaria *R* explica una acción *A* sólo en relación con una descripción de esa acción. Por ejemplo, suponiendo que mi acción de hacer girar el conmutador sea idéntica a mi acción de dar la alarma por la presencia de un ladrón, la acción no está explicada (racionalizada) por mi deseo de encender la luz (ni por mi creencia de que hacer girar el conmutador tendrá por efecto encender la luz) sino como acción de hacer girar el conmutador. Pero lo que es válido en el caso de las razones (o los efectos) lo es generalmente también en el caso de las acciones (o las causas); si mi deseo de encender la luz es idéntico a mi deseo de realizar la acción designada en inglés con la expresión *turning on the light* (o a tal estado neurofisiológico), solamente como deseo de encender la luz explica (racionaliza) mi acción (en cuanto a acción de hacer girar el conmutador). Esto muestra que habiendo una racionalización de la acción *A* por la razón *R* por medio de descripciones “*A*” y “*R*”, el resultado de la sustitución de otra descripción de *A* o de *R* no constituye siempre una racionalización ni una explicación causal. Por consiguiente, las características mentales son causalmente pertinentes. Esto sugiere que las condiciones impuestas al concepto de racionalización primaria deben reformularse a fin de que sean sensibles no sólo a las descripciones de las acciones sino también a las de las razones primarias. Se podría decir entonces que la razón primaria  $R = (x, y)$  en la descripción “*R*” es una razón por la que un agente ha realizado la acción *A* en la descripción “*A*” solamente si:

- 14) “*R*” es una descripción de *x* en cuanto proactitud del

agente respecto de acciones que satisfacen cierta descripción  $d$  y una descripción de  $y$  como creencia del agente de que  $A$  en la descripción "A" satisface la descripción  $d$ .

15)  $R$  causa  $A$

y

16) "R" y "A" son causalmente pertinentes.

Las condiciones 15) y 16) equivalen a la condición de que  $R$  explica causalmente  $A$ . Estas tres condiciones son evidentemente más vigorosas que aquellas a que antes aludimos, pero luego mostraré que así y todo no son todavía suficientes.

Todo esto es perfectamente compatible con la negación de CNTCP, es decir, con el hecho de que partiendo de la circunstancia de que un evento mental  $m$  explique causalmente (o racionalice) un evento físico  $p$ , no se puede llegar a la conclusión de que  $m$  y  $p$  ejemplifiquen una ley si satisfacen respectivamente las descripciones " $m$ " y " $p$ ". Davidson dispone pues aparentemente de una respuesta inmediata a la objeción de Honderich. Pero lo importante aquí es observar que el hecho de que las descripciones utilizadas dentro del marco de una racionalización sean causalmente pertinentes no implica que tales descripciones sean las únicas causalmente pertinentes de los mismos sucesos. Por el contrario, en virtud de CNC, debe haber también descripciones (y éstas no pueden ser sino físicas) de los mismos eventos que están en una relación no sólo explicativa sino también nomológica. Manifiéstase pues que lo que (en esta perspectiva) hace que dos descripciones estén en una relación explicativa es el hecho de que ellas ejemplifiquen un esquema explicativo "reconocido", a saber, tanto el esquema de la explicación nomológica como el de la racionalización (o, más generalmente, el esquema de la explicación causal singular). Para sostener que las características mentales son causalmente pertinentes, Davidson debería pues sostener por lo visto que la racionalización constituye un modo de explicación causal singular, no sólo legítimo sino también irreductible.<sup>5</sup>

Stoutland y Honderich tienen tendencia a confundir sus respectivas críticas (Stoutland 1985: 53; Honderich 1983: 148). Esas críticas, en efecto, se confunden fácilmente si se supone, como parece ocurrir en el caso de Stoutland y de Honderich, que toda explicación causal debe ser nomológica; sólo que Stoutland prefiere expresarse con el vocabulario de la explicación antes

que con el de las relaciones nomológicas, de modo que importa verificar en qué medida la pertinencia de sus críticas depende de este presupuesto.

En realidad, Stoutland señala varias dificultades en el análisis de Davidson y todas las hace derivar del hecho de que ese análisis no permite explicar por qué un evento mental dado es idéntico a un determinado suceso físico (1980: 357; 1985: 51). Aparentemente Stoutland cree reformular la misma crítica diciendo que no puede haber (en el análisis de Davidson) relación explicativa entre el hecho de que un evento sea un evento mental y el hecho de que ese evento posea estos o aquellos poderes causales (1980: 357). Sin embargo estas dos formulaciones sólo son equivalentes si hay que admitir que un evento mental sólo tiene poder causal en virtud de sus características físicas, admisión que precisamente acabo de impugnar. Pero si se suprime esta relación entre las dos formulaciones no es claro que la primera constituya una objeción seria. Si el evento mental  $m$  es idéntico al evento físico  $p$ , aparentemente éste debe ser un hecho inexplicable porque no puede deberse ni a la existencia de una equivalencia nomológica entre una característica mental de  $m$  y una característica física de  $p$ , ni a la existencia de una relación causal entre  $m$  y  $p$ . Pero considerar esto como una objeción equivale sencillamente a declarar indefendible la tesis de la *token*-identidad. Debemos pues atender a la segunda formulación que es, por lo demás, aquella que invoca Stoutland (1980: 357-58) para concluir que en el análisis de Davidson no hay relación explicativa entre la razón mencionada por una racionalización de una acción y esa acción misma. También esta vez la conclusión es correcta sólo presuponiendo que no son las características físicas de las razones las que son causalmente pertinentes. Si en cambio se admite la pertinencia causal de las características mentales es trivialmente cierto que no hay relación explicativa entre el hecho de que un evento tenga tales características y el hecho de que tenga tales poderes causales (puesto que se trata de una relación de identidad), pero no se puede llegar a la conclusión de que las razones no expliquen las acciones.

Esta interpretación no hace del todo justicia a Stoutland. Este podría sin duda replicar que no niega que haya una relación explicativa entre las razones y las acciones sino que niega que sea posible explicar por qué una razón  $R$  causa una



acción dada o distinguir de manera general entre las ocasiones en las que una razón R causa una acción de cierto tipo y las ocasiones en que dicha razón no causa ninguna acción de ese tipo (1980: 363; 1985: 51-52). Este último punto pone particularmente de relieve lo que, según creo, es la cuestión principal de esta controversia, es decir, la cuestión de establecer si es posible en principio predecir la conducta de un sujeto sobre la base del conocimiento de sus características psicológicas.

El principal motivo para sugerir que la existencia de una relación causal entre una razón R y una acción A sea una condición necesaria es, en efecto, que esto permite distinguir entre las ocasiones en que un sujeto tiene la razón R sin cumplir la acción A por esa razón y aquellas en que cumple la acción A por esa razón. Stoutland pretende que esta condición no permite hacer la distinción deseada porque en general no se sabe distinguir entre las ocasiones en que R causa A y las ocasiones en que R no causa A. Si esto significa que Stoutland pide que sean precisadas las condiciones empíricas que permitan reconocer los casos en que R racionaliza A, entonces parece lícito responder que la sugestión de Davidson sólo tiene un alcance ontológico y no apunta a cumplir esa función, sino que pretende tan sólo precisar en qué consiste (en parte) la diferencia entre los casos en que R racionaliza A y los demás casos.

Pero cualesquiera que sean las intenciones de Davidson, puede parecer legítimo exigir que se precisen las condiciones empíricas generales que permitan justificar de manera concluyente la afirmación de una relación de causalidad (o de explicación causal) singular. Entonces la objeción de Stoutland se reduce a esto. Davidson admite que es necesario para afirmar, por ejemplo que la razón R causa la acción A, apoyarse en generalizaciones psicológicas relativamente bien confirmadas que relacionan por ejemplo tipos de creencia o de deseo, o de ambos, y tipos de conducta. Sin embargo esas generalizaciones tienen la índole de leyes estrictas (especialmente porque están invariablemente sometidas a condiciones *ceteris paribus*) y por eso nunca pueden constituir una justificación completa de la afirmación de que R causa A. Si en cambio R y A son descripciones físicas relacionadas por una ley estricta (como deben estarlo, según Davidson), el conocimiento de esa ley (y de esas descripciones) permite afirmar con certeza que R causa A y por consiguiente también explicar o predecir A sobre la base de una

descripción física de R. Hasta aquí no hay ninguna razón para alarmarse. La dificultad planteada por Stoutland procede tan sólo del hecho de que, como no hay explicación de que R admita esta o aquella descripción física, tampoco puede haber explicación nomológica de que R *explique causalmente* A (es decir, explicación del hecho de que R causa A *en cuanto razón*). Aquí Stoutland parece presuponer que toda explicación causal debe ser nomológica y reprochar a Davidson que admita una concepción de la racionalización en virtud de la cual es imposible predecir la conducta de un agente sobre la base de un conocimiento de sus estados intencionales. Interpretada de esta manera, la objeción de Stoutland no se distingue de la de Honderich y su suerte depende de la suerte de la tesis según la cual la racionalización es un modo de explicación causal legítimo e irreductible aunque desprovisto de valor predictivo. Si aquí desgraciadamente me es imposible hacer el examen de las consideraciones pertinentes para evaluar esta tesis, la discusión anterior habrá por lo menos permitido hacer resaltar su importancia para apreciar justamente la naturaleza de la diferencia que separa a Davidson de Stoutland y Honderich.

Pero me falta todavía mencionar otro aspecto de la crítica de Stoutland, que también puede interpretarse como la exigencia de que se especifiquen condiciones *suficientes* para que una razón racionalice una acción (y no sólo para que ella la explique causalmente). Stoutland (1980: 365; 1985: 53) observa, en efecto, que la concepción davidsoniana de la racionalización hace insoluble el problema de las cadenas causales atípicas. Pero si es cierto que Davidson (1980: 79, 87, 232-233) considera insoluble este problema, esto no se debe, según me parece, al hecho de que una razón no puede, según él, explicar una acción en cuanto razón pues el problema permanece igualmente insoluble en la hipótesis contraria.

Ya he puesto de manifiesto que las dos condiciones impuestas por Davidson para que una razón racionalice una acción no eran suficientes. El motivo de ello es el de que es perfectamente posible que un agente tenga una razón primaria R para hacer A y que R cause efectivamente A sin que podamos decir intuitivamente que el agente ha hecho A por la razón R. Davidson (1980: 79) da el siguiente ejemplo: un alpinista puede desear librarse del peso y del peligro que constituye para él la circunstancia de sostener a su compañero por el extremo de una

cuerda y saber que si suelta la cuerda podría liberarse de ese peso y de ese peligro. Esta razón primaria puede provocar en el alpinista un estado de nerviosidad tal que afloje la cuerda. Manifiestamente, ese agente tiene una razón para soltar la cuerda y esa razón es la causa de haberla aflojado, pero no afloja la cuerda *por* esa razón; en realidad su acción no es en modo alguno intencional. Para excluir este tipo de casos Davidson agrega que no basta con que una razón primaria cause una acción para que la racionalice; todavía hace falta que la cause *de manera correcta* (*in the right kind of way*). El problema de las cadenas causales atípicas se concibe entonces como el problema de precisar, sin circularidad, de qué manera una razón debe causar una acción para poder racionalizarla. En otras palabras, se trata de decir en qué condiciones una acción resulta de un razonamiento práctico del agente utilizando únicamente conceptos como los de creencia, de proactitud o de causalidad. Davidson (1980: 232-233) considera irrealizable esta tarea, pues no ve cómo se podría dar cuenta de las normas de racionalidad que entran en el concepto de razonamiento práctico sin explicar cómo un agente evalúa la fuerza relativa de deseos o de creencias mutuamente incompatibles, es decir, sin recurrir a nociones vedadas (como por ejemplo la de indicación (evidencia) o la de “una buena razón para creer”).

Stoutland observa (1985: 53), refiriéndose al ejemplo de Davidson que acabo de mencionar, que el problema se plantea porque la conducta del agente no está causada por su razón *en cuanto razón*, sino que está causada por su razón en cuanto fuente de perturbación emocional y que Davidson no puede resolver el problema porque no admite que una razón puede causar una acción *en cuanto razón*. Esto suscita tres comentarios. En primer lugar, si es cierto que Davidson nunca dice que un evento causa otro *en esta o aquella descripción*, he sugerido antes que podría tener sentido esta manera de hablar y que ella corresponde a la manera que tiene Davidson de comprender la relación de explicación causal. La expresión “en cuanto razón” puede entenderse de dos maneras en el contexto que nos ocupa. A saber, en el sentido de “en cuanto a tener esta o aquella característica mental” o en el sentido de “en cuanto elemento de un razonamiento práctico”. Ahora bien, en segundo lugar, es claro que, si uno interpreta la expresión de la segunda manera y la utiliza para enunciar las condiciones suficientes para que

una razón racionalice una acción (suponiendo que esto sea posible), se violará entonces la prohibición de no circularidad impuesta por Davidson. Y en tercer lugar, cualquiera que sea la interpretación escogida, se puede concebir perfectamente que en el ejemplo mencionado es en cuanto razón como las actitudes del agente provocan la nerviosidad (por lo demás, ésta me parece la interpretación más natural del ejemplo); sin embargo no se sigue de ello que la conducta del agente esté racionalizada por esa razón, lo cual muestra que el problema no queda suprimido admitiendo la pertinencia causal de las características mentales (y que sin duda tampoco quedará suprimido admitiendo la pertinencia causal de las características “racionales”). En realidad, me parece que el razonamiento práctico de un sujeto puede ser tan complejo y explícito como se quiera sin por ello causar la acción pertinente *de la manera correcta*, por más que la cause *en cuanto razonamiento práctico*, pues siempre queda abierta la posibilidad de que cause una reacción emotiva intermedia. Y quizá por esta razón Davidson duda de que sea posible especificar el tipo de proceso causal apropiado para la racionalización de la acción, aun de manera circular.

Concluyo examinando una posible objeción a la idea de que admitir la pertinencia causal de las características mentales constituya una respuesta suficiente a la acusación de epifenomenismo. Esa objeción está hecha por Honderich (1982: 64) cuando interpreta la convicción de que las características mentales son causalmente pertinentes como si fuera la convicción de que éstas son una parte ineliminable de toda explicación completa de una acción. Ahora bien, parecería que Davidson no pudiera admitir la pertinencia causal de las características mentales si se lo interpreta de esta manera. Según la concepción de Davidson, en efecto, parece que toda acción que pueda explicarse causalmente por una razón puede también (en virtud del carácter nomológico de la causalidad) explicarse invocando una ley estricta que relacione una descripción física de esa razón con una descripción física de esa acción. De suerte que la racionalización se manifiesta entonces como una explicación redundante.

Pero como lo indican Macdonald y Macdonald (1986: 44), este razonamiento es falaz porque no tiene en cuenta la intencionalidad de la relación explicativa. Del hecho de que un evento mental en la descripción *m* explique causalmente un

hecho (físico o mental) en la descripción  $p$  y del hecho de que exista una ley estricta que relacione una descripción física  $x$  de  $m$  con una descripción física  $y$  de  $p$ , no se puede en modo alguno llegar a la conclusión de que  $m$  en la descripción  $x$  explique causalmente  $p$  en la descripción  $y$ . Para ello sería menester que hubiera una relación explicativa (que en este caso no podría ser sino nomológica) entre  $y$  y  $p$  (o que  $y = p$ ); ahora bien, en el caso en que  $p$  es una descripción mental esto queda excluido por el anomalismo de lo mental, en tanto que en el caso en que  $p$  es una descripción física la existencia de semejante relación no está de ninguna manera garantizada (puesto que todas las descripciones físicas de un mismo hecho no están nomológicamente relacionadas). Por consiguiente, el hecho de que  $m$  satisfaga la descripción mental " $m$ " puede perfectamente ser un elemento esencial de toda explicación de  $p$  que satisfaga la descripción " $p$ " (que se puede, por ejemplo, suponer como una descripción tomada del lenguaje ordinario de la acción).

Es interesante observar que esta respuesta no depende de la tesis del anomalismo de lo mental como tal, sino que depende de una hipótesis a la vez más débil y más general, a saber, la hipótesis de que puede no haber ninguna relación nomológica entre dos descripciones " $y$ " y " $p$ " de un mismo evento si éste está en una relación de explicación causal en la descripción " $p$ " (por ejemplo) con un evento mental en la descripción mental " $m$ " y si hay una relación nomológica entre " $y$ " y una descripción física " $x$ " de  $m$ . En el caso contrario, por cierto no se ve cómo se podría negar que hay una relación explicativa entre " $p$ " y " $x$ ", lo que quiere decir que la descripción " $m$ " es superflua, esto es, o bien que es posible explicar  $p$  en la descripción " $p$ " sin hacer intervenir la descripción " $m$ " (en el caso en que  $m$  causa  $p$ ) o bien que  $p$  en la descripción " $p$ " no puede explicar  $m$  en la descripción " $m$ " sin explicarla también en la descripción " $x$ " (en el caso en que  $p$  causa  $m$ ). Esta hipótesis me parece suficientemente plausible para que sea lícito concluir, por lo menos provisionalmente, que el monismo anomal continúa siendo una opción perfectamente viable.<sup>6</sup>

## Notas

<sup>1</sup> Davidson propuso primero este principio como criterio de identidad en el caso de los eventos. Si recientemente (1985: 175) renunció a utilizarlo de esta manera no parece que se pueda llegar a la conclusión de que Davidson lo considere en adelante falso.

<sup>2</sup> Esta crítica parece también estrechamente relacionada con una dificultad planteada por Blackburn (1985) en relación con el predominio psicofísico. Llegado el momento oportuno espero poder explicar esta relación.

<sup>3</sup> Honderich prefiere hablar de propiedades antes que de rasgos o de descripciones. Pero eso no tiene ninguna importancia a los efectos de la presente discusión.

<sup>4</sup> Macdonald y Macdonald (1986: 43) llegan por otro camino a lo que parece ser esencialmente la misma conclusión.

<sup>5</sup> Tal vez hasta bastaría con sostener que la racionalización es una relación no extensional y que las razones son causas, sin pronunciarse sobre el valor "explicativo" de la racionalización, valor que de todas maneras es ciertamente menor que el de la explicación nomológica.

<sup>6</sup> El contenido de este artículo fue objeto de dos exposiciones presentadas el 30 de abril de 1987 en la universidad de Québec en Trois Rivières y el 4 de junio de 1987 en el seminario "Filosofía del espíritu" (CNRS). Quiero agradecer a S. Davis, P. Engel, N. Kaufman, A. Leclerc, C. Panaccio, J. Proust, C. Savary, M. Seymour y R. Vallée por sus comentarios y sus palabras de aliento.

## Referencias bibliográficas

- Apel, Karl-Otto (1981) "The Erklaren-Verstehen Controversy in the Philosophy of the Natural and Human Sciences"; Floistad, G. (comp.) (1981-83), vol. 2, 19-50.
- Armstrong, D.- M. (1983) "Recent Work on the Relation of Mind and Brain"; Floistad, G. (comp.) (1981-83), vol. 4, 45-80.
- Bacon, J. (1986) "Supervenience, Necessary, Coextension and Reducibility", *Phil. Studies* 49, 163-176.
- Biro, J. - I. y R.-W. Shahan (comps.) (1982) *Mind, Brain and Function*; Norman, Univ. of Oklahoma Press.
- Blackburn, S. (1985) "Supervenience Revisited"; Hacking, I. (comp.) (1985), 47-57.
- "Finding Psychology"; Stevenson, L. et al. (comps.) (1986), 1-12.
- Block, N. (comp.) (1980) *Readings in the Philosophy of Psychology*; vol. 1, Londres, Methuen.
- Borgen, R. y F. Cioffi (comps.) (1970) *Explanation in the Behavioral Sciences*; Cambridge, Cambridge Univ. Press.
- Boyd, R. (1980) "Materialism without Reductionism: What Physicalism does not Entail"; Block N. (comp.) (1980), 67-106.
- Currie, G. (1984) "Individualism and Global Supervenience"; *Brit. J. for the Phil. of Sc.* 35, 345-358.
- Davidson, D. (1980 a) *Essays on Actions and Events*; Oxford, Oxford Univ. Press.
- (1980 b) "Toward a Unified Theory of Meaning and Action"; *Grazer Phil. Studien* 2, 1-12.
- (1982) "Rational Animals"; Lepore, E. y B. McLaughlin (comps.) (1985), 473-481.

(1984 a) *Inquiries into Truth and Interpretation*; Oxford, Oxford Univ. Press [Hay versión castellana: *De la verdad y de la interpretación*. Barcelona, Gedisa, 1990.]

(1984 b) "First-Person Authority"; *Dialectica* 38, 100-111.

(1985 a) "Replies to Essays I-IX"; Vermazen, B. y M.-B. Hintikka (comps.) (1985), 195-229.

(1985 b) "Replies to Essays I-XII"; Vermazen, B. y M.-B. Hintikka (comps.) (1985), 242-252.

(1985 c) "Reply to Quine on Events"; Lepore, E. y B. McLaughlin (comps.) (1985), 311-323.

(1985 d) "Adverbs of Action"; Vermazen, B. y M.-B. Hintikka (comps.) (1985), 230-241.

Davies, M. (1986) "Externality, Psychological Explanation and Narrow Content", *Proc. of the Arist. Soc. Supp.* vol. 60, 263-283.

Elgin, C. (1980) "Indeterminacy, Undetermination and the Anomalism of the Mental"; *Synthese* 45, 233-255.

Elster, J. (1981) "Rationality"; Floistad, G. (comp.) (1981-83), vol. 2, 111-132.

Engel, P. (1986 a) "L'anomalie du mental"; *Critique* n° 474, 1125-1139.

(1986 b) "Le mental dépend-il du physique?", inédito.

Feldman, F. (1980) "Identity, Necessity and Events", Block, N. (comp.) (1980), 148-155.

Floistad, G. (comp.) (1981-83) *Contemporary Philosophy*, 4 vols, Dordrecht, Martinus Nijhoff.

Fodor, J. (1986) "Individualism and Supervenience"; *Proc. of the Arist. Soc. Supp.* vol 60, 235-262.

Follesdal, D. (1985) "Causation and Explanation: A Problem in Davison's View on Action and Mind"; Lepore, E. y B. McLaughlin (comps.) (1985), 311-323

French, P.-A. et al (comps.) (1979) *Midwest Studies in Philosophy 4: Studies in Metaphysics*; Minneapolis, Univ. of Minnesota Press.

French, P.-A. et al. (comps.) (1984), *Midwest Studies in Philosophy 9: Causation and Causal Theories*, Minneapolis, Univ. of Minnesota Press.

Hacking, I. (comp.) (1985), *Exercises in Analysis*; Cambridge, Cambridge Univ. Press.

Hare, R.-M. (1984) "Supervenience"; *Proc. of the Arist. Soc. Supp.* vol. 58, 1-16.

Haugeland, J. (1982) "Weak Supervenience", *American Phil. Quart.* 19, 93-103.

(1984 a) "Ontological Supervenience", *Southern J. of Phil.* 22, Supp., 1-12.

(1984 b) "Phenomenal Causes", *Southern J. of Phil.* 22, Supp. 71-92.

Hellman, G. y F.-W. Thompson (1975) "Physicalism: Ontology, Determination and Reduction", *J. of Phil.* 72, 551-567.

(1977) "Physicalist Materialism", *Nouÿs*, 11, 309-345.

Honderich, T. (1982) "The argument for Anomalous Monism", *Analysis* 42, 59-64.

(1983) "Anomalous Monism: Reply to Smith", *Analysis* 43, 147-149.

(1984) "Smith and the Champion of Mauve", *Analysis* 44, 86-89.

Horgan, T. (1981) « Token Physicalism, Supervenience and the Generality of Physics », *Synthese* 49, 395-413.

(1984) "Supervenience and Cosmic Hermeneutics", *Southern J. of Phil.* 22, Supp., 19-38.

Horgan, T. y M. Tye (1985) "Against the Token Identity Theory"; Lepore, E. y B. McLaughlin (comps.) (1985), 427-443.

Horsby, J. (1980) "Wich Physical Events are Mental Events?", *Proc. of the Arist. Soc.* 81, 73-92.

(1985) "Physicalism, Events and Part-Whole Relations"; Lepore, E. y B. McLaughlin (comps.) (1985), 444-458.

Jeffrey, R. (1985) "Animal Interpretation"; Lepore, E. y B. McLaughlin (comps.) (1985), 408-426.

Johnston, M. (1985) "Why having a Mind Matters"; Lepore, E. y B. MacLaughlin (comps.) (1985), 408-426.

Kernohan, A. (1985) "Psychology: Antonomous or Anomalous?", *Dialogue* 24, 427-442.

Kim, J. (1979) "Causality: Identity and Supervenience in the Mind-Body Problem"; Franch P. A. et al (comps.). (1979), 31-50.

(1978) "Supervenience and Nomological Incommensurables", *America Phil. Quart* 15, 149-156.

(1982) "Psychophysical Supervenience", *Phil Studies* 41, 51-70.

(1984 a) "Epiphenomenal and Supervenient Causation"; French P. A. et al (comps.) (1984), 257-270.

(1984 b) "Supervenience and Supervenient Causation"; *Southern J. of Phil.* 22, Supp., 45-56.

(1984 c) "Self-Understanding and Rationalizing Explanations", *Philosophia Naturalis* 21, 309-320.

(1984 d) "Concepts of Supervenience", *Phil. and Phenom. Research* 45, 153-176.

(1985) "Psychophysical Laws"; Lepore E. y B. McLaughlin (comps.) (1985), 369-386.

Lepore, E. (1985) "The Semantics of Action, Event and Singular Causal Sentences"; Lepore E. y B. McLaughlin (comps.) (1985), 151-161.

Lepore, E. y B. McLaughlin (comps.) (1985) *Actions and Events*, Oxford, Blackwell.

Lepore, E. y B. McLaughlin (1985) "Actions, Reasons, Causes and Intentions", Lepore, E. y B. McLaughlin (comps.) (1985), 3-13.

Levinson, J. (1984) "A Thousand Entities", *Southern J. of Phil.* 22, Supp., 13-18.

Lewis, H.-A. (1985) "Is the Mental Supervenient on the Physical?"; Vermazen B. y M.-B. Hintikka (comps.) (1985), 159-172.

Lycan, W.-G. (1982) "Psychological Laws"; Biro, J.-I. y R.-W. Shahan (comps.) (1982), 9-38.

Macdonald, G.C. Macdonald (1986) "Mental Causes and the Explanation of Action, Stevenson L. et al (comps.) (1986), 35-48.

Mackie, J.-L. (1979) "Mind, Brain and Causation"; French, P.-A. et al (comps.) (1979), 19-30.

McDowell, J. (1985) "Functionalism and Anomalous Monism"; Lepore E. y B. McLaughlin (comps.) (1985), 387-398.

McGinn, C. (1977) "Anomalous Monism and Kripke's Cartesian Intuitions"; Block, N. (comp.) (1980), 156-158.

(1978) "Mental States, Natural Kinds and Psychophysical Laws", *Proc. of the Aust. Soc. Supp.* vol. 52, 195-220.

(1980) "Philosophical Materialism"; *Synthese* 44, 173-206.

McLaughlin, B. (1984) "Event Supervenience and Supervenient Causation", *Southern J. of Phil.* 22, Supp. 71-92.

(1985) "Anomalous Monism and the Irreducibility of the Mental"; Lepore, E. y B. McLaughlin (comps.) (1985), 331-368.

Melchert, N. (1986) "What's Wrong with Anomalous Monism?", *J. of Phil* 83, 265-274.

Millikan, R.-G. (1986) "Thoughts without laws", *Phil. Rev.* 95, 47-80.

Mohanty, J.-N. (1984) "Intentionality, Causality and Holism"; *Synthese* 61, 17-33.

Nelson, R.-J. (1976) "Mechanism, Functionalism and the Identity Theory", *J. of Phil* 73, 365-385.

Peters, R.-S. (1970) "Comment"; Borger R. y F. Cioffi (comps.) (1970), 27-48.

Post, J.-F. (1984) "Comment on Teller", *Southern J. of Phil.* 22, Supp. 163-167.

Putnam, H. (1983 a) *Realism and Reason*, Cambridge, Cambridge Univ. Press.

(1983 b) "Computational Psychology and Interpretation"; Putnam, H. (1983 a), 139-154.

Quine, W.-V.O. (1985) "Events and Reification"; Lepore, E. y B. McLaughlin (comps.) (1985), 162-171.

Robinson, H. (1982) *Matter and Sense*, Cambridge, Cambridge Univ. Press.

Rorty, R. (1982) "Contemporary Philosophy of Mind", *Synthese* 53, 323-348.

Rosenberg, A. (1985) "Davidson's Unitended Attack on Psychology", Lepore, E. y B. McLaughlin (comps.) (1985), 399-407.

Seager, W.-E. (1981) "The Anomalousness of the Mental", *Southern J. of Phil.* 19, 389-401.

Smart, J.-J.-C. (1985) "Davidson's Minimal Materialism"; Vermazen, B. y M.-B. Hintikka (comps.) (1985), 173-182.

Smith, P. (1982) "Bad News of Anomalous Monism?", *Analysis* 42, 220-224.

(1984) "Anomalous Monism and Epiphenomenalism: A Reply to Honderich", *Analysis* 44, 83-86.

Sosa, E. (1984) "Mind-Body Interaction and Supervenient Causation"; French, P.-A. et al (comps.) (1984), 271-282.

Stanton, W.-L. (1983) "Supervenience and Psychophysical Law in Anomalous Monism", *Pacific Phil Quart.* 64, 72-79.

Stevenson, L. et al (comps.) (1986) *Mind Causation adn Action*, Oxford, Blackwell.

Stich, S. (1978) "Autonomous Psychology and the Belief-Desire Thesis", *The Monist* 61, 573-591.

Stoutland, F. (1982) "Philosophy of action: Davidson, von Wright and the Debate over Causation"; Floistad, G. (comp.) (1981-83), vol. 3, 45-72.

(1980) "Oblique Causation and Reasons for Action ", *Synthese* 43, 351-367.

(1985) "Davidson's on Intentional Behavior"; Lepore, E. y B. McLaughlin (comps.) (1985), 44-59.

Suppes, P. (1985) "Davidson's Views on Psychology as a Science"; Vermazen, B. y M.-B. Hintikka (comps.) (1985), 183-194.

Teller, P. (1984a) "A Poor Man's Guide to Supervenience and Determination", *Southern J. of Phil.* 22, Supp. 137-182.

(1984 b) "Comments on Kim's Paper"; *Southern J. of Phil* 22, Supp., 57-62.

Toulmin, S. (1970) "Reasons and Causes"; Borger, R. y F. Cioffi (comps.) (1970), 1-26.

Tuomela, R. (1982) "Explanation of Action"; Floistad, G. (comp.) (1981-83), vol. 3, 15-44.

Tye, M. (1984) "Supervenience, Materialism and Functionalism", *Southern J. of Phil.* 22, Supp. 39-44.

Vermazen, B. y M.-F. Hintikka (comps.) (1985) *Essays on Davidson: Actions and Events*, Oxford, Clarendon Press.

Wilkes, K.-V. (1982) "Functionalism, Psychology and the Philosophy of Mind "; Biro, J.-I. y R.-W. Shahan (comps.) (1982), 147-167.

De la colección de PAPELES JPG  
en <http://padron.entretemas.com.ve>

## Un dilema para el fisicismo <sup>1</sup>

### Criterios estructurales y funcionales de la identidad de los estados mentales\*

Howard Robinson

#### 1

El fin de este artículo es mostrar que hay ciertas circunstancias en las que un fisicista con vocación funcionalista debe tratar la existencia de cierto rol conceptual como algo que define un tipo particular de estado mental y que hay circunstancias en las que ese fisicista se ve obligado a tratar la existencia de ciertos rasgos estructurales, caracterizados bioquímicamente o neurofisiológicamente y separados de su función, como elementos suficientes de la constitución de estados mentales; y me propongo mostrar además que no se pueden conciliar estos dos requisitos de una manera que no sea arbitraria.

La forma más ingenua del fisicismo es la que puede llamarse "teoría pura de la identidad". Se trata de la teoría de que un análisis filosófico de los estados mentales de cierto tipo M no necesita más que esos estados bioquímicos o neurofisiológicos, sin referencia esencial alguna a la función de tales estados en la regulación de la conducta. La idea principal, por supuesto es, no que los B causan M o que los M surjan de algún modo por relación con los B, sino que los M son sencillamente el estado fisiológico pertinente, así como el agua es sencillamente H<sub>2</sub>O. Ningún fisicista contemporáneo sostiene esta teoría, que yo sepa. Todos ellos están de acuerdo en afirmar que la naturaleza fisiológica de B *per se* no explica aquello que

\* En la edición original francesa este artículo fue traducido del inglés por P. Engel [E.]

hace que los B son M y que es necesario explicar por qué los B constituyen M. Es en este punto donde el fisicista se vuelve hacia el funcionalismo al invocar el papel causal de los B para explicar por qué un trozo de materia de estructura B podría contarse, por ejemplo, como un dolor. A partir del momento en que se aparta uno de la pura teoría de la identidad para encaminarse hacia el funcionalismo puede o bien hacer un análisis mixto de los estados mentales que incluyen condiciones a la vez estructurales y funcionales, o bien dirigirse por entero hacia un análisis funcional que descarta la referencia a la estructura. Sin embargo, en esta última interpretación el funcionalismo puro fracasa desde el comienzo. "El funcionalismo puro" es la teoría según la cual hallarse en un estado mental M consiste en hallarse en un estado que ejerce la función apropiada F, es decir, que la función debe ser actualizada. Que esta teoría fracase o no depende de la cuestión de si se consideran realizaciones centrales o realizaciones totales de los estados funcionales. La realización central es, en términos generales, el estado pertinente del sistema nervioso central y la realización total es el conjunto de los mecanismos requeridos para conectar estímulo y respuesta, incluso los nervios periféricos y los músculos. Si el funcionalismo quisiera insistir en el hecho de que estos últimos estén actualizados para que el sujeto se encuentre en el estado central apropiado, sostendría una teoría tan poco plausible como la forma más ingenua del conductismo y excluiría del dominio de los estados mentales a) los casos en los que no se siguiera ninguna conducta y b) los casos en que el estímulo llegara directamente al cerebro por un medio consecuente, ¡a primera vista la única forma posible de funcionalismo puro!

Este artículo se propone mostrar que una teoría mixta y el funcionalismo central son falsos.

#### 2

Consideraré primero diversas teorías mixtas. La más común es la propuesta por David Lewis.<sup>2</sup> Reformulando ligeramente sus términos iniciales la posición de Lewis puede expresarse así:

A) S siente un dolor si (y solamente si) S se encuentra en el estado que tiene ordinariamente la función del dolor en el tipo de especie a que pertenece S.

Lewis deriva esta teoría de la necesidad de dar cuenta del “dolor loco”, es decir, el caso supuesto y posible de un ser humano en el que el estado del cerebro (que en el resto de nosotros es causado por una lesión y determina una conducta de dolor) está causado por estímulos completamente diferentes con efectos también completamente diferentes. Ese hombre sentirá un dolor cuyas relaciones causales son enteramente atípicas. Pero la posibilidad del dolor loco presenta serios problemas.

El problema para el loco del que habla Lewis es el siguiente. Lewis insiste en el hecho de que su análisis es un análisis del dolor tal como lo comprendemos normalmente, en otras palabras, es un análisis del carácter fenoménico de la experiencia, del *efecto que hace (what it is like)* ser un loco que estuviera en ese estado y de *cómo* el loco *experimenta ese estado*. Es decir, no se puede dar nuevo sentido al “dolor” que sea independiente del sentido que damos nosotros a esa palabra en nuestra experiencia.<sup>3</sup> Pero supongamos que, por más que el loco sea lo que es, el resto del género humano sea diferente aunque su fisiología sea la misma que la del loco. El dolor loco del “loco” ya no será ese estado que en los otros miembros de nuestra especie desempeña la función causal propia del dolor y, por lo tanto, sentirá un dolor cuando se encuentre en el estado “loco”. Parece raro decir que (aunque para el fisicista tener un dolor sea una condición puramente física) la estructura física de un individuo no sea por sí sola suficiente para determinar si ese individuo siente un dolor.<sup>4</sup>

Se puede mostrar con mayor claridad que la posición de Lewis no es plausible si consideramos el caso siguiente. Supongamos que se construya una nueva clase de robot dotado de sensaciones, lo cual debe ser posible de conformidad con el tipo de teorías del espíritu que estamos examinando. Llamémoslo un robot Mark V. Hasta ahora sólo se ha construido un Mark V y se lo ha llamado Joe y Joe es un Mark V loco en el sentido siguiente: por una razón cualquiera, el estado que debería haber tenido, según los planos, las funciones correspondientes al estado normal de dolor —llamémoslo el estado X— tiene otra función. ¿Siente dolor Joe cuando se encuentra en el estado X? Se podría responder “sí” diciendo que X es el estado que, según los planos, funciona como dolor en ese modelo de robot. Pero supongamos que los constructores, sintiéndose enteramente satisfechos con la conducta de Joe, decidan construir todos los otros robots Mark V como Joe. Hemos de considerar los casos siguientes:

I) Todos los Mark V sienten dolor cuando se encuentran en X, aunque ninguno de ellos se conduce como si experimentara dolor.

II) Una vez contruidos los demás Mark V, X dejará de ser un estado de dolor porque éste ya no es el estado normal de dolor para esa especie; de manera que Joe no sentirá ya dolor en el estado X.

III) Ninguno de los demás Mark V experimentará dolor pero, porque Joe lo sentía en el caso inicial, ¿continuará sintiéndolo en este caso?

Ninguna de estas opciones es aceptable. I) No es aceptable porque algo que estaba inscripto en el plano, pero nunca fue actualizado, seguramente no tiene nada que ver con las experiencias reales de una criatura capaz de sensaciones. Supongamos que esto haya sido incluido en el plano por error o por una decisión arbitraria; o supongamos que un minuto antes de ser construido Joe el plano hubiera sido modificado de manera que X ya no fuera funcionalmente el estado de dolor ordinario o que se lo hubiera modificado durante la construcción de Joe o un minuto después de ésta. Es completamente ridículo tratar estos hechos como no pertinentes en lo relativo a la cuestión de saber lo que experimentan efectivamente los robots Mark V. II) Es inaceptable porque hace arbitraria la experiencia de Joe. Supongamos que éste padezca un dolor crónico, es decir que X se declare continuamente. Si aceptamos II), mientras se construyen los otros robots, Joe, aunque tal vez completamente inconsciente de las estructuras internas de estos aparatos, deja de experimentar dolor por más que X continúe declarándose en él y por más que físicamente nada cambie en él, simplemente porque X ya no es dolor en el caso de Mark V. III) Supone abandonar el principio mismo que admite el dolor loco, pues si Joe puede tener una experiencia en virtud de un estado cervical diferente del estado que constituye esa experiencia para los otros miembros de su especie, ¿por qué los otros Mark V no pueden tenerla? El hecho de que Joe sea el original parece arbitrario y no pertinente. Pero si admitimos esta clase de desviaciones del principio según el cual la existencia del estado funcional ordinario constituye el dolor, ya no tenemos ninguna razón para afirmar que el “dolor loco” es realmente dolor. Se puede defender III) diciendo que Joe no es un miembro de la misma especie que los otros Mark V,

pero esta estipulación sería arbitraria si Joe satisficiera todos los demás criterios a los que responde Mark V, especialmente en la medida en que todo Mark V de la serie, semejante a Joe, será de la misma especie que los demás y simplemente un robot loco.

La otra posibilidad consiste en negar que Joe originalmente sintiera dolor cuando se declaraba X a pesar del plano que asignaba a X la función de dolor en el caso de los Mark V. La razón de ello sería la de que un simple plano no puede determinar lo que constituye la norma de la especie. Pero de todas maneras se trata de dilemas inaceptables. Supongamos que los demás Mark V de la serie estén contruidos con estados X cuya función es el dolor, ¿cuál es entonces la experiencia de Joe y cuál es la de los demás Mark V? Aquí hay tres opciones:

a) X es ahora dolor para todos los Mark V, incluso Joe, por más que el dolor de Joe sea dolor loco que le causa esa sensación pero que no le provoca una conducta de dolor;

b) X es dolor para los otros, pero no para Joe;

c) ninguno de ellos siente dolor cuando está en X porque la norma de la especie estaba establecida por Joe en su condición de miembro original de la especie, de manera que todos los demás Mark V son locos que se conducen como si sintieran dolor siendo así que no lo sienten.

Ninguna de estas opciones es aceptable. a) Es inaceptable por la misma razón que II), a saber, la experiencia de Joe se considera modificada simplemente a causa de cambios externos que no tienen ningún efecto físico sobre él. b) Tampoco es pertinente por razones semejantes a III), a saber, si admitimos que un miembro de la especie siente dolor en un estado estructural diferente del estado del resto de la especie, la explicación que está en la base de la teoría de Lewis se desmorona. c) No es en modo alguno plausible. Los Mark V pueden permanecer sin cambiar como especies o como individuos durante millones de años; pueden llegar a ser la única forma de ser consciente de la galaxia. c) Requeriría que, a causa de la manera en que ha sido hecho Joe, esa manera no afectara al resto de los robots, que estos no experimentaran dolor ni siquiera en el momento mismo en que se condujeran como si lo experimentaran; aunque, si nunca se hubiera construido Joe, el resto de los robots aun permaneciendo físicamente sin cambios habría experimentado dolor.

Hemos mostrado que A) es inadecuado y que por lo menos debe ser reemplazado por

B) S siente un dolor si S se encuentra en el estado que ordinariamente tiene la función de S en el interior de S mismo.

La conclusión sugerida por la discusión de Lewis es la de que en el caso del dolor y, por analogía de razonamiento, en el caso de otras sensaciones, el estado mental prevalece (*supervene*) sobre la constitución física del organismo. Conceptualmente no se requiere ningún otro hecho de tipo social o ambiental; lo que se requiere será sólo causalmente para producir el estado interno en cuestión. La idea de que nada fuera del organismo pueda ser pertinente al estado de sensación que ese organismo experimenta, puede llamarse principio del carácter intrínseco. El criterio A) de Lewis no logra satisfacer ese principio, no sólo porque hace que la cuestión de saber si una criatura tiene una sensación depende de la manera en que están contruidos *otros* objetos, sino también porque (según lo muestran los argumentos que hemos expuesto) Lewis admite diferentes decisiones sobre el punto de establecer qué experiencia se experimenta y porque vacía de contenido la tesis de que una cualquiera de estas interpretaciones pueda ser más verdadera que otra. Hasta B) es sin embargo insuficientemente precisa en cuanto al principio del carácter intrínseco.

Supongamos que, tal como está ahora construido S, el estado sea el estado mental que posee la función de dolor. Y supongamos que, por obra de la neurocirugía o por cualquier otra transformación de la organización neuronal de S, Y quede privado de esa función que se le da en cambio a Z. Y es entonces estimulado. ¿Experimenta S dolor? Una criatura —llamémosla M— que estuviera construida originalmente como S, *después del cambio*, no experimentaría dolor cuando su estado Y fuera estimulado según B). Pero S después del cambio (digamos sólo S para abreviar) es físicamente semejante a M con toda exactitud. La naturaleza misma del principio del carácter intrínseco sugiere que S y M deberían estar en el mismo tipo de estado de sensaciones. Verdad es que nuestra refutación de Lewis no justifica directamente un principio de carácter intrínseco tan fuerte como éste, pero justificaría que se excluyera la referencia a cualquier cosa exterior a S, aunque no la referencia a sus propios estados internos. Pero en la medida en que S y M son exactamente semejantes en cuanto a lo físico y en la medida en



que la memoria aparente predomina sobre el estado físico actual (ésta es una hipótesis que suponemos que un fisicista quiere hacer), por más que el pasado de S tenga actualmente influencia en él, esa influencia estará representada en su estado presente. Intuitivamente, si la operación quirúrgica alterara todas las relaciones funcionales de Y y de Z, incluso las de la memoria, parece que lo que funcionalmente es dolor para S estaría alterado por la operación. S no tiene necesidad de saber nada sobre el cambio sobrevenido y no experimentará nada diferente, de modo que el nuevo estado será su estado de dolor. Tenemos necesidad, no de B), sino de

C) S siente un dolor si se encuentra en el estado que (en él tal como está ahora construido) tiene la función del dolor.

En este estadio, la distinción entre las realizaciones centrales y totales de las funciones se hace pertinente. Por una razón que ya hemos mencionado, esto es, que los estados mentales pueden sobrevenir en ausencia del proceso funcional total, las entidades candidatas a la identidad con los estados mentales se consideran en general estados del sistema nervioso central (SNC), en tanto que la función estímulo-respuesta es realizada por una red mucho más importante de músculos y de nervios. Hemos de ver que la importancia de estos estados en el caso de las funciones constituye un grave problema para el fisicista, sea que este decida incluirlos en su análisis o no. Si no los incluye, esto puede expresarse explícitamente reemplazando C) por

D) S siente un dolor si se encuentra en el estado central que, en su SNC tal como está actualmente constituido, tiene el papel funcional del dolor.

Para comprender por qué D) es inadecuado, debemos presentar a un nuevo personaje, VC.

### 3

“VC” es la abreviatura de “víctima del curare”. Pero hay dos variantes de este personaje y sólo una de ellas es exactamente semejante a una verdadera víctima del curare. Los rasgos comunes a las dos versiones son que VC queda enteramente paralizado y no puede mostrar ninguna reacción a nada. En el segundo caso, que está idealizado, VC queda igualmente cortado del contacto sensorial normal con el mundo, de manera

que las experiencias puedan ser causadas por estímulos aplicados directamente al sistema nervioso central.

Mantendré la hipótesis de que nadie puede dudar de que el VC normal pueda experimentar dolor, etc. Recientemente algunas madres de familia del Lancashire recibieron indemnizaciones por daños y perjuicios porque dieron a luz mediante una operación cesárea estando totalmente anestesiadas por el curare pero no anestesiadas correctamente, pues las mujeres sintieron toda la operación o parte de ella aunque no pudieron manifestar su dolor en ese momento. Supongo que nadie duda de que alguien que se encontrara en el lugar de esas madres experimentaría un gran dolor en el momento de la operación. Acaso la hipótesis de que el sistema aferente y el sistema deferente estén “cortados”, como en el caso de una VC idealizada, podría hacer más problemática la tesis de que la experiencia puede sobrevenir incluso cuando el SNC es directamente estimulado. Un wittgensteiniano, por ejemplo, podrá sostener que en ese caso ninguno de los criterios para atribuir una experiencia (se refiera, ya al estímulo, ya a la respuesta) es satisfactorio. La pertinencia del enfoque wittgensteiniano de esta cuestión se manifiesta a medida que se va desarrollando la argumentación, pero por el momento haré sólo la siguiente hipótesis. Supongamos que una víctima (en oposición a una víctima idealizada) del curare sea inducida a experimentar dolor por estímulos normales de dolores en el cuerpo. Su sistema aferente también queda fuera de circuito y recibe un estímulo equivalente en esos centros receptores del dolor. Cuando el sujeto recobra la conciencia comunica que ha experimentado dos dolores exactamente semejantes. Supondré que no queremos decir que el primer dolor era experimentalmente real y que el segundo no lo era. En otras palabras, supongo que nuestras intuiciones sobre VC normales se aplican también al caso idealizado, y a los efectos de la simplicidad entenderé por “VC”, en lo que sigue, el caso idealizado más extremo, salvo que dé aviso de lo contrario. El aparente carácter plausible de casos como las dos variantes de VC es una de las razones por las que los materialistas han abandonado el conductismo y se han vuelto hacia la teoría de la identidad espíritu-cuerpo; porque en efecto VC tiene los estados del cerebro apropiados al dolor si sus centros receptores del dolor son directamente estimulados, aun cuando VC no manifeste ningún signo de reacción en su conducta. El problema de

toda teoría que intenta hacer esos estados del cerebro conceptualmente y no sólo causalmente suficientes para explicar los dolores es el problema de establecer en virtud de qué semejante estado ha de considerarse como dolor. Esta es la dificultad que ya señalamos al hablar de la teoría pura de la identidad y es una dificultad que reconocen los físicos contemporáneos serios: el problema está en que nada en la estructura química intrínseca (haciendo abstracción de sus propiedades causales) puede llamarse un dolor, de modo que hay que hacer un análisis cualquiera de las razones por las que ese proceso debería llamarse dolor. El único análisis plausible (y éste es el que se hace en general) es un análisis que atiende al papel causal de esas razones: semejantes estados son habitualmente producidos por una lesión del organismo y habitualmente provocan lo que llamamos “una conducta de dolor”. Pero en el caso de VC, el estado del cerebro apropiado no está causado por una lesión, ni da lugar a la respuesta apropiada. Supongamos que alguien diga que cuando falta el papel funcional ya no se trata de dolor; ¿qué puede decir el materialista contra esta sugestión?

Al llegar a este punto el materialista puede recurrir a una teoría mixta de la cual el único ejemplo que sobrevive es D), es decir, el materialista puede sostener que el proceso cervical constituye el estado mental, porque es el estado estructural que asegura corrientemente la función apropiada en este organismo particular VC. En primer lugar pudiera parecer que hay otra opción, a saber, adoptar el funcionalismo central puro. Según esta última teoría, el estado del cerebro se cuenta como un estado mental, no a causa de la función de conducta que él asegura *ordinariamente*, sino a causa de su actual función en el interior del sistema nervioso central. La dificultad que presenta esta última teoría (como se verá con mayor claridad luego) es que el SNC en sí mismo no tiene salida *de conducta* del tipo de las salidas que son pertinentes en el caso de la atribución de los estados mentales. Estados del SNC son mentales, no en virtud de las salidas electroquímicas del SNC, sino en virtud del papel de esas salidas en la producción de la conducta. Esa es su función normal y es lo que los hace estados mentales, según los criterios funcionalistas, de modo que la segunda posibilidad aparente no difiere de la primera, presuponiendo su función ordinaria en su conjunto.<sup>5</sup>

La finalidad de la mayor parte del resto de este artículo es sostener que el físico no puede aceptar que VC haya sufrido realmente en el momento en que sus centros receptores de dolor fueron activados (artificialmente o naturalmente) sino que el físico se ve obligado a aceptar la interpretación “malcolmiana” de la situación de VC. Se podrá precisar esta sugestión así. Supongamos que alguien que haya leído el libro de Malcolm\* sobre los sueños sugiera que el curare aumentado con un estímulo del cerebro es la causa de que alguien haya tenido recuerdos falsos y la causa de que ese alguien comunique incorrectamente su experiencia al cesar el efecto de la droga. ¿Donde está la falsedad de esta interpretación “malcolmiana”? Esta tesis no se formula como una duda escéptica. En otras palabras, no se basa en alguna ontología física cualquiera, pues nada puede constituir la verdad de la concepción del sentido común sobre VC contra la concepción “malcolmiana”. Los únicos hechos físicos pertinentes son que VC ha sido drogado, que sus centros receptores han sido estimulados, que hasta el momento VC no ha manifestado ninguna respuesta y que al cesar el efecto de la droga se queja vigorosamente de que ha experimentado dolor. La cuestión de establecer si realmente experimentó dolor en el momento del estímulo no es algo que dentro de esta ontología pueda determinarse como hecho establecido (*fact of the matter*). La teoría “malcolmiana” no niega ninguno de los hechos físicos. El problema para el físico que piensa que VC ha experimentado realmente dolor (o que las madres del Lancashire tuvieron una compensación legítima) está en decir (en sus propios términos) en qué el “malcolmiano” se equivoca, puesto que el malcolmiano no niega ningún hecho físico.

En primer lugar, el físico podría responder que una víctima real del curare, a diferencia del soñante de Malcolm, puede mostrarse consciente de otras maneras que cuando es víctima de la influencia de la droga. Por ejemplo, si se le muestran diversos objetos cuando se halla bajo esa influencia, en el momento de salir de ella el sujeto será capaz de comunicar lo que ha visto. Por consiguiente es consciente y por lo tanto habrá experimentado el dolor. Pero uno puede evitar esta respuesta o bien afinando lo que el malcolmiano pudiera decir

\* Norman Malcolm, *Dreaming*, Londres, Routledge, 1962.

o bien alterando ligeramente el ejemplo. Se podría insistir en que si bien VC tiene realmente conciencia de los objetos, recuerda equivocadamente haber experimentado dolor, o en que VC no era consciente de nada en aquel momento, pues solamente habría recibido la información cuando estaba en trance, como quien estuviera sujeto a una sugestión hipnótica, lo cual le permite comunicar luego lo que vio. O bien uno podría alterar un poco el ejemplo y proponer la forma reforzada de VC a que antes aludimos e imaginar que la droga —u otros anestésicos locales— corta todos los contactos sensoriales con el mundo exterior, pero dejando activos los estados centrales. No dudaremos de la aptitud modificada de VC para experimentar dolor, para pensar, etc. En todos estos casos, el “malcolmiano” puede insistir, sin estar en desacuerdo con los procesos físicos en cuestión, en que VC no siente nada mientras dura la operación.

Un no “malcolmiano” podría insistir en que esta situación no representa para él ningún problema porque el “malcolmiano” comete un error conceptual. D) es una verdad analítica y VC se encuentra en el estado funcional apropiado, por más que sus funciones estén inhibidas, de manera que el hecho de que esa persona está sufriendo es un hecho analítico. Expondré dos argumentos separados contra esta objeción de la tesis “malcolmiana”. El primero está destinado a mostrar que ella minimiza la cuestión de establecer si VC y las madres del Lancashire sufren realmente. El segundo está destinado a demostrar que VC no se encuentra en el estado funcional apropiado.

#### 4

*La tesis de que el error malcolmiano es puramente conceptual minimiza el problema.* Supongamos que alguien sugiera que, si bien (de conformidad con nuestro concepto actual de dolor) VC realmente sufrió, deberíamos modificar ese concepto de manera tal que dijéramos que el dolor no existió sino cuando se produjo una reacción efectiva cualquiera. Por las razones que ya hemos dado, un físico no puede rechazar esta sugestión alegando que es empíricamente errónea, pues los hechos físicos están fuera de toda discusión. A veces ocurre que uno pueda rechazar sugerencias de revisión conceptual sobre la base de que nuestros actuales conceptos están demasiado bien afirma-

dos para que pueda realizarse en la práctica la revisión; pero la sugestión propuesta se refiere a un elemento tan marginal de nuestra práctica —la parálisis total es rara— que no se podría hacer este tipo de objeción. Se podría más bien sostener que no hay alteración porque las personas tratadas como VC continuarían quejándose de haber sufrido un dolor. Pero podemos modificar nuestra réplica a esto y en lugar de indemnizarlas por daños y perjuicios, procurar tratar sus recuerdos como si uno procurara liberar a alguien de una pesadilla que se repite frecuentemente. Hasta podríamos adoptar como nuestro “anestésico” algo que paralice y que elimine el recuerdo del dolor en el momento de despertar. Suponiendo que por cualquier razón adoptar esta medida pudiera resultar cómodo, un físico podría objetarla únicamente por razones que tuvieran que ver con su posición conservadora conceptual. Sin embargo, no creemos que sea respondiendo a la simple posición conservadora conceptual que se diga que la existencia de dolores agudísimos en el momento de la operación es un hecho bruto, de manera que una revisión conceptual que ignorara esto ignoraría algo fundamental. Como la revisión “malcolmiana” no ignora ninguno de los hechos físicos, parece que el hecho del cual creemos que ella ignora debe de ser no físico, según los criterios físicos reduccionistas.

#### 5

*VC no se encuentra en el estado funcional apropiado.* Demostrar esta conclusión implica demostrar dos proposiciones: I) Ningún estado central (estado del sistema nervioso central) es suficiente por sí solo (es decir, haciendo abstracción de su localización) para constituir un estado mental particular de conformidad con criterios funcionales; II) pero si otros estados se incluyen en el estado funcional, VC no satisface los criterios que definen el estado funcional de dolor (o, por supuesto, ningún otro estado psicológico), salvo en un sentido “malcolmiano”, en el mejor de los casos. Se puede demostrar I) por el siguiente argumento:

1) Cualquier máquina puede tener papeles funcionales significativamente diferentes según el sistema global a que pertenece.

2) El SNC es una máquina.

3) Todo SNC puede tener diferentes papeles funcionales según el sistema global al que pertenece.

4) La identidad de los estados del SNC como (*en cuanto*) estados mentales depende de la significación funcional que éstos tengan para el organismo particular en que están localizados.

Por consiguiente:

5) Todo estado del SNC puede tener identidades mentales diferentes según el organismo particular en que está localizado.

Considero la primera premisa como una perogrullada que se debe a la manera en que se entienda el concepto de "máquina", a saber, todo aquello que funciona según principios físicos en general deterministas. Por ejemplo, un determinado motor de combustión tiene una salida funcional diferente si está adaptado a un automóvil o si funciona como motor fijo industrial en un garaje o si se lo adapta a una bomba para desagotar una mina. Hasta el motor del automóvil completo podría estar en el interior de una máquina mayor cuyas salidas nada tuvieran que ver con la locomoción, máquina en la que sin embargo contribuirían las diferentes funciones del automóvil.

2) Es también una perogrullada para el fisicista. Por consiguiente: 3) se sigue por instanciación universal y *modus ponens*. El hecho de que las diferentes disposiciones posibles de un SNC sea un asunto de complejidad propio de la ficción científica nada tiene que ver aquí. 4) Puede ser aceptado por quienquiera que piense que los estados del cerebro no pueden considerarse estados mentales, haciendo total abstracción del papel que ellos desempeñan en la producción final de la conducta; en otras palabras, alguien que niega que sea puramente contingente la circunstancia de que los estados mentales produzcan una conducta observable. Todo partidario de una teoría causal que se sintiera tentado por semejante teoría debería quedar disuadido por 3), pues haciendo abstracción de sus localizaciones, los estados propios del SNC en su conjunto tienen una identidad mental indefinidamente disyuntiva. 5) Se sigue de 3) y 4) atendiendo al carácter transitivo de la relación de dependencia; en otras palabras, atendiendo al hecho de que, si la identidad mental depende de la función y si la función depende de la naturaleza del sistema en el cual está localizada, luego, la identidad mental depende de la localización.

Para defender II) —la tesis de que no podemos admitir la pertinencia de la localización en el caso de VC— hay que mostrar que I) sólo se aplica a VC al precio de consecuencias inaceptables. Es natural negar que I) no permite a VC decir que experimenta dolor porque el SNC de VC está localizado en otro sistema, cuyo funcionamiento normal debería conducir a la conducta de dolor. Pero, por otra parte, el curare hace que VC se parezca a alguien que no tiene semejante sistema: quíerase o no, durante el tiempo en que está paralizado, VC es una criatura que carece de ese otro sistema. Supongamos que en el momento mismo en que quedó paralizado pero justo antes de aplicarse el estímulo a sus centros receptores de dolor se hayan cortado las conexiones entre el SNC de VC y el resto de su cuerpo. Sus centros receptores estarían entonces directamente estimulados y el SNC estaría desconectado. No dudamos de que cuando hubieran pasado los efectos del curare VC se quejaría de haber experimentado dolor como si la conexión entre su SNC y el resto de su cuerpo nunca hubiera sido cortada; tampoco dudamos de que VC se quejaría de manera perfectamente justificada.

Consideremos los casos siguientes:

1) VC es víctima del curare ordinario, de manera que puede estimulárselo normalmente, aunque no pueda reaccionar. Su dolor está producido por una conexión física corriente y cuando recobra la conciencia VC pretende haber sufrido.

2) VC se encuentra ahora en el caso crítico con los sistemas aferentes y deferentes paralizados. Se estimulan sus centros receptores de dolor y cuando recobra la conciencia VC pretende haber sufrido.

3) VC está paralizado, pues las conexiones entre su SNC y el resto de su cuerpo están cortadas; se estimulan sus centros receptores, se vuelve a conectar el SNC y al recobrar la conciencia VC dice que recuerda haber sufrido.

4) VC está paralizado; su SNC es enteramente extirpado de su cuerpo, se estimulan los centros receptores, se inserta el SNC en su cuerpo y se lo reconecta. Al recobrar la conciencia VC pretende recordar que ha sufrido.

5) Como 4), sólo que después de la aplicación del estímulo el SNC se inserta en un sistema que da un valor funcional enteramente diferente a una parte del SNC o a todo el SNC, incluso los centros receptores y los recuerdos asociados; de

manera que VC no pretende tener el recuerdo de haber sufrido pero tiene una reacción diferente.

6) Un SNC, elaborado *in vitro* idéntico a VC célula por célula con activación de lo que en CV representa el papel de receptor, nunca fue insertado en un organismo cualquiera.

La demostración de I) prueba que el SNC de 6) no tiene los estados funcionales para que se pueda decir que posee una psicología. En el momento de aplicarse el estímulo, el SNC de VC en 4) y 5) es exactamente similar (en cuanto a lo físico) al SNC de 6) y por consiguiente, y según el principio del carácter intrínseco, se encuentra en ese momento exactamente en los mismos estados mentales que en 6), a saber, en ningún estado mental particular. La diferencia entre 3) y 4) es simplemente la dramática aparición de la extracción del SNC; aquí lo que cuenta es la desconexión. Así el SNC de VC en 3) y en el momento del estímulo es también como el SNC de 6). Si se considera lo que VC experimenta (*what it is like*), no puede haber diferencia entre 2) y 3) pues la droga “congela” totalmente la relación de VC con sus estados periféricos, de suerte que para él es como si éstos no existieran. Por consiguiente VC en 2) no puede tener ninguna de las sensaciones que VC no tiene en 3). El SNC es pues en 2) como el SNC de 6); VC no tiene psicología. Ya he sostenido al presentar el VC normal y el VC idealizado que en principio 1) no era diferente de 2) de modo que 1) no es diferente de 6) desde el punto de vista de la experiencia realizada. Cuando cotejamos el hecho de que el SNC paralizado no tiene psicología con el hecho de que hay una diferencia entre 5) y 1), 2), 3), 4), podemos comprender por qué la interpretación “malcolmiana” es la interpretación correcta para el físico.

El estímulo no tiene valor mental específico mientras el SNC no esté alojado funcionalmente en alguna parte; si se lo reimplanta en un nuevo contexto como en 5), puede no tener el valor de un estado de dolor, pero si se lo reintegra en otro ser humano normal, tendrá ese valor. Esto no se podía establecer en aquel momento, pero se presentará como un hecho de esos que son apropiados en el caso de la localización dentro de un organismo cuando se restauran las relaciones funcionales. Esto corresponde exactamente a la teoría “malcolmiana” de la situación de VC.

Hemos discutido y considerado inadecuadas todas las teorías que intentan explicar el estado de conciencia de VC (en

ausencia de un funcionamiento real) haciendo referencia a estados dotados de una estructura bioquímica o neurofisiológica apropiada. Lo que acabo de decir incluye a D), la única versión que quedaría sin refutar de la teoría mixta funcional y estructural.

## 6

La cuestión de saber si estas conclusiones representan una amenaza seria para el fisicismo resulta clara si evaluamos la situación. Sobre esta situación hay dos diagnósticos que pueden rechazarse fácilmente.

I) X se encuentra en un estado de dolor cuando dicho estado sobreviene en S si X tiene cierta composición física

o

II) X se halla en un estado de dolor si este sobreviene en S y si X ejerce simultáneamente el papel funcional explícito del dolor.

Hemos rechazado categóricamente I) al mostrar que no nos explica por qué un estado físico de ese tipo deba atribuirse a una entidad física. II) no es pertinente porque no puede dar cuenta de casos en los que la función no se ejerce. Todos los físicos tratan de combinar I) y II) para elaborar algo de la forma general de

III) X se halla en un estado de dolor cuando este sobreviene en S y S está en un estado de tipo físico que ejerce de manera ordinaria la función del dolor.

Las opciones de A) a D) que he considerado entran todas en esta última categoría; difieren sólo en cuanto a lo que determina lo que haya de entenderse por “ordinario”. La refutación de las tesis de A) a D) constituye la refutación de III), que pareciera ser ella misma la última esperanza para el físico. Sin embargo habrá físicos que rechacen todo intento de hacer un análisis de los conceptos psicológicos en términos compatibles con el fisicismo: la “psicología ordinaria” es sencillamente falsa en ciertos aspectos y no puede ser enteramente compatible con la verdad física. Sospecho que sentimientos tales están en la base de observaciones como la de Wilkes:

“La cuestión de si algo debe ser clasificado como una descripción ‘funcional’ o ‘estructural’ es con frecuencia

arbitraria y sin interés... La distinción es mucho menos importante de lo que suponen numerosos funcionalistas.”<sup>6</sup>

Comentarios de este género reflejan la impaciencia que siente el materialista eliminador frente a la “psicología ordinaria” que es, según ellos suponen, precientífica y falsa. Lo que espera hacer el partidario eliminador es mostrar cómo las explicaciones físicas son lo único valioso del trabajo de la “psicología ordinaria”. La “psicología ordinaria”, después de habernos puesto en la pista de los mecanismos neuronales al llevarnos a examinar ciertas conexiones funcionales interesantes, se borra y nosotros nos interesamos por las estructuras cervicales en sí mismas; éstas no nos interesan ya como portadoras de funciones que las hagan corresponder a nuestros conceptos precientíficos de la “psicología ordinaria”. De esta manera, el materialista eliminador reivindica la ventaja de la teoría de la identidad pura (es decir, que se puede prescindir de la noción de función) sin sentirse obligado a hacer que esa identidad corresponda a los hechos referentes a la psicología usual.

El caso de VC es ideal para mostrar la impropiedad de este intento de eludir el problema. Como sostuve cuando traté la cuestión de establecer si un cambio conceptual podría legítimamente afectar lo que debíamos pensar de la experiencia de VC, lo que realmente se plantea es el problema de saber si VC estaba sometido o no a horribles tormentos cuando se hallaba bajo los efectos de la droga. Si lo estaba, sus sufrimientos no son sólo parte de una teoría —destinada a ser abandonada pronto o por lo menos destinada a ser abandonada en principio—, pues toda nueva teoría debe sostener por lo menos como un hecho establecido que VC sufrió realmente durante todo el tiempo. Si el materialista eliminador quiere satisfacer esta condición, se ve obligado a replegarse a la teoría de la identidad pura y a decir que todo lo que era correcto cuando hablábamos del dolor en este caso incumbía a los procesos del cerebro. Pero entonces el materialista debe darnos un análisis de la razón por la cual esos procesos del cerebro deben interesarnos, y hacerlo sin volver a introducir la psicología ordinaria; de otra manera no podría explicar por qué el estado en que se encontraba VC era tan lastimoso ni por qué las madres del Lancashire debían recibir

indemnización por daños y perjuicios. Si ese materialista quiere mantener la creencia de que VC y aquellas madres sufren realmente, nos debe una explicación dada en sus propios términos. Parece que el materialismo eliminador no puede permitir al fisicismo eludir el dilema según el cual ni criterios funcionales y estructurales ni una combinación particular de esos criterios pueden suministrar una explicación adecuada de por qué ciertos estados físicos son también estados conscientes, porque el dolor de VC es solamente el tipo de fenómeno del cual pensamos que sería francamente falso “eliminarlo”.

## 7

¿Adónde conducen al fisicista los anteriores argumentos?

La mayor parte de lo que precede se basa en la premisa implícita de que cualquiera experimenta dolor en la clase de circunstancias que he considerado y de que éste es un hecho crudo, de manera que para un fisicista este hecho debe prevalecer (*supervene*) sistemáticamente en los estados físicos. Esto se puede llamar realismo fisicista en lo referente al dolor y ese realismo resultaba explícitamente en lo que he llamado el principio del carácter intrínseco. Hasta los comentarios sobre el materialismo eliminador pueden considerarse a la luz de este marco realista en la medida en que el materialismo eliminador considera la existencia del dolor como un hecho crudo *supuesto*, pero un hecho que sencillamente no existe. Lo que muestra el fracaso de todas las teorías fisicistas que he examinado es que resulta imposible ofrecer un análisis del dolor que sea a la vez fisicista y realista. Y lo que esto quiere decir exactamente puede comprenderse si se examinan las siguientes posiciones. Las siguientes tesis fueron consideradas y rechazadas.

a) El SNC de un sujeto que esté en el estado, por así decirlo, normal X es suficiente sencillamente —es decir, sin mecanismos suplementarios— para que se dé el dolor.

b) El SNC de un sujeto que se encuentre en un estado interno X es suficiente para que se dé el dolor. a) y b) quedaron definitivamente rechazadas por dos argumentos importantes del párrafo 5 respectivamente. Todo cuanto queda es la idea de que factores enteramente exteriores al sistema pueden ser conceptualmente pertinentes. En efecto, esto querría decir que

las formas sociales y lingüísticas, formas de vida y juegos de lenguaje, son esenciales para que algo constituya un dolor. Sin embargo este requisito es por lo menos ambiguo, pues deja sin decidir la cuestión de establecer si, en el seno de nuestras formas sociales y lingüísticas, el dolor prevalece en ciertos estados internos. En otras palabras, se podría adoptar una especie de realismo interno como en:

c) Atendiendo a nuestras formas sociales y lingüísticas, el hecho de que el SNC de un sujeto esté en un estado interno X es suficiente para que se dé el dolor.

Esto contrasta con el no realismo interno de que

d) El hecho de que el SNC de un sujeto esté en cierto estado interno sea suficiente para que se dé el dolor depende de factores sociales y lingüísticos que pueden conducir a decisiones diferentes en diversos casos en los que los sujetos tienen esos SNC en X, es decir, cuando no hay diferencia descollante en el SNC mismo.

En d) se pierde el predominio de todo lo que es central y queda enteramente abandonado el principio del carácter intrínseco en el caso de las sensaciones. Los argumentos de este artículo obligan al fisicista a aceptar d) porque van contra c), contra a) y contra b). Además, el argumento del párrafo 4 —que sostiene que hacer del dolor del VC una cuestión de convención minimiza el problema —es en realidad un argumento contra c) en particular. Ahora bien, lo que muestran los argumentos de este artículo es que no existen convenciones fundadas en el concepto de estado central que sean coherentes y aceptables y que rijan la existencia de dolores, lo cual era la tesis requerida por c) y el realismo fisicista habitual. Es posible pues presumir que el fisicismo debe retirarse a lo que se puede llamar un instrumentalismo en cuanto a las sensaciones, un instrumentalismo paralelo al que Dennett recomienda en el caso de los estados intencionales. Saber si una u otra de estas posiciones puede constituir una oposición fundada al enfoque “malcolmiano” de VC es un punto no aclarado, por más que el argumento expuesto en el párrafo 4 sugiera que ninguna de esas posiciones puede serlo. Tal vez el fisicista debiera abandonar la intuición y aceptar la interpretación “malcolmiana” de VC. En este artículo mi intención no era perseguir al fisicista en estas regiones oscuras y dudosas, sino que era mostrar que en ellas se ve obligado a residir el fisicista.

## Notas

<sup>1</sup> Una versión de este artículo fue dada como conferencia en la serie “Filosofía del espíritu” con los auspicios del Centro Nacional de Investigación Científica (CNRS), París, 1986. Esa versión fue asimismo leída ante grupos de estudiosos de Liverpool y de Oxford, y agradezco a las numerosas personas que me brindaron su ayuda en el curso de esos encuentros.

<sup>2</sup> D. Lewis, “Mad Pain and Martian Pain”, en *Readings in the Philosophy of Psychology*, vol I, Londres, Methuen, 1980, págs. 216-22.

<sup>3</sup> *Ibid.*, págs. 211-22.

<sup>4</sup> Sobre una observación análoga, véase Sidney Shoemaker “Some varieties of functionalism” en *Identity, cause and mind*, Cambridge, Cambridge University Press, 1984, págs. 261-86; esta referencia está en la pág. 269.

<sup>5</sup> No es raro que los fisicistas ignoren la distinción entre propiedades disposicionales inmediatas de los procesos del SNC y disposiciones de largo plazo, productoras de conducta. Semejante elisión tiene una parte importante en la versión de la teoría de la identidad espíritu-cerebro de D.M. Armstrong; véase H. Robinson, *Matter and sense*, Cambridge, 1982, págs. 52-58.

<sup>6</sup> K.V. Wilkes, “Is consciousness important?”, *British Journal for the Philosophy of Science*, vol. 35, 1984, págs. 222-43.

## Glosario

**Actitudes proposicionales.** Los estados intencionales tales como la creencia, el deseo, el anhelo, etc. cuyos contenidos “pro-posicionales” son expresados lingüísticamente por oraciones completas regidas por *que* (por ejemplo, Juan cree que *p*).

**Análogo/digital.** Según uno de los sentidos de esta distinción, un sistema de representación es digital si está compuesto de símbolos sintácticos desunidos, diferenciados de manera acabada y discreta, y análogo es el sistema denso y continuo. Según otro sentido de la distinción, un proceso es análogo si su dimensión es una magnitud física (considérese la distinción entre ordenadores análogos y digitales).

**Condición de generalidad.** Según G. Evans, una condición que rige toda posesión de conceptos y de pensamientos y, según la cual, no se puede afirmar que un sujeto posea un concepto si no está en condiciones de formar juicios que impliquen a otros sujetos de esos conceptos, o de relacionar sujetos de juicios con otros conceptos.

**Conexionismo.** En las ciencias cognitivas una corriente actual que, a diferencia del cognitivismo “clásico”, trata las computaciones no como operando en sistemas simbólicos gobernados por reglas, sino como operando en redes asociativas paralelas y distribuidas (“PDP”) que, según se supone, representan un modelo del funcionamiento de las redes neuronales.

**Conocimiento tácito.** En principio, el conocimiento no consciente almacenado en los estados cognitivos, por ejemplo, aquellos de los que se supone que son responsables de la capacidad de hablar una lengua. Se plantea la cuestión de saber si es un conocimiento en el sentido usual.

**Contenido estrecho/amplio.** Según una terminología debida a H. Putnam, el contenido de un pensamiento o de una

creencia es estrecho si sólo apela a estados psicológicos interiores e individuales del sujeto y es amplio si hace referencia a entidades o a propiedades semánticas especificables al mencionar las condiciones de verdad de los contenidos y el ambiente del sujeto.

**Encapsulamiento informacional.** Según Fodor, propiedad de los sistemas modulares de tratamiento de la información, propiedad que los opone a las creencias y a los pensamientos de los sistemas centrales: la información de los sistemas modulares es aislada y propia de una determinada tarea y no puede volver a utilizarse en otros contextos ni para otros fines.

**Entrada-salida** (*input/output*) en la jerga de la teoría de la información: entrada-salida de información en un sistema.

**Epifenomenismo.** En metafísica tradicional, tesis de que lo mental no tiene ninguna eficacia causal, en particular sobre los sucesos físicos.

**Externalismo.** Tesis según la cual los contenidos de los estados mentales intencionales están determinados y caracterizados por propiedades del ambiente exterior de los sujetos.

**Funcionalismo.** En filosofía del espíritu, tesis según la cual los estados mentales tienen papeles causales y funcionales en sistemas de entrada-salida de información, como en el caso de los estados del programa de un ordenador. Consecuencia: los estados mentales pueden realizarse en cualquier composición material (*hardware*).

**Identidad (teoría de la).** En filosofía del espíritu, versión de la tesis materialista de que los eventos mentales son idénticos a eventos físicos, ya genéricamente (*tipos*), ya como hechos particulares (*tokens*).

**Individualismo:** Concepción según la cual la naturaleza y carácter de los contenidos intencionales dependen exclusivamente de la psicología interna e individual de un sujeto y no del ambiente, por oposición al externalismo.

**Modularidad.** Hay varios conceptos de modularidad, pero el de Fodor es el invocado con mayor frecuencia. Intuitivamente la modularidad caracteriza los sistemas cognitivos que son autónomos y están aislados dentro de la arquitectura del espíritu. Un sistema cognitivo es modular si es específico de un dominio (por ejemplo, la visión, el tratamiento de la información fonética), si es inherente a él, si está relacio-



nado con un sistema neuronal localizado y estructurado, si es computacionalmente autónomo y produce una información “encapsulada” y no reunida, es decir, no construida partiendo de subprocesos elementales. Los sistemas periféricos son modulares, por oposición al sistema central de la cognición.

**Monismo anomal.** Teoría de lo mental debida a D. Davidson, formada por la conjunción, para él coherente, de tres tesis aparentemente incompatibles: 1) los eventos mentales tienen interacciones con los físicos, 2) donde hay causalidad hay ley y 3) no hay leyes deterministas estrictas de lo mental. Se trata de una forma de materialismo mínimo que identifica hechos mentales y hechos físicos (monismo) y que, por otro lado, rechaza la idea de leyes de lo mental y de leyes psicofísicas (anomal).

**PDP-Parallel Distributed Processing.** Según algunos, religión reciente practicada sobre todo en la costa occidental de los Estados Unidos. Véase *conexionismo*.

**Posición eliminatoria.** En filosofía del espíritu, forma de materialismo según la cual el vocabulario y los estados mentales deben, no reducirse sino eliminarse en provecho de un vocabulario solamente neurofisiológico. Tesis de que asimismo la psicología ordinaria debe eliminarse.

**Prevalecer (supervenience).** Inicialmente, término de la filosofía moral que designa una relación de dependencia y determinación de un grupo de propiedades por otro grupo. Una (de las) propiedad (es) P prevalece o predomina sobre una (de las) propiedad (es) R si es imposible que dos objetos tengan en común todas sus propiedades P, pero difieran por sus propiedades R.

**Principio del carácter intrínseco.** Véase *solipsismo metodológico*.

**Sistemas periférico/central (input/central systems).** En la terminología de Fodor, sistemas (modulares) que aseguran las entradas de información del mundo exterior y corresponden generalmente a una modalidad sensorial o a un tratamiento de información aislado.

**Solipsismo metodológico.** Principio metodológico formulado por Putnam, según el cual la psicología cognitiva debe atenerse a una explicación de los contenidos de estados mentales puramente internos e individuales del sujeto, con exclusión de sus relaciones con el ambiente. Es el principio metodológico correspondiente (pero no idéntico) a la doctrina individualista.

**Subdoxasticidad.** Caracteriza todos los estados representacionales internos de la cognición situados en un nivel infrapersonal y a menudo inconsciente; está en la base de los estados conscientes de actitudes proposicionales (por ejemplo, los estados comprendidos en el proceso de la visión en las fases iniciales).

**Token/tipo.** Distinción (inicialmente debida a Peirce y a la semiótica) entre las apariciones concretas y particulares de un signo y la categoría genérica del signo (por ejemplo, en la vida cotidiana pronunciamos numerosos *tokens* del tipo “palabrota”). Por extensión, caso particular versus caso genérico.

**Transductor.** En electrónica, dispositivo que recibe flujos de energía y los retransmite en una forma alterada. En ciencias cognitivas, dispositivo que permite pasar del nivel físico y material de un sistema al nivel simbólico/computacional. Por ejemplo, el tipo de procesos que está en la base del “esbozo primario” en la teoría de la visión de D. Marr que toma propiedades de los estímulos visuales para disponerlos en informaciones métricas sobre, por ejemplo, la distancia de los objetos.

**Transmodular.** Proceso cognitivo común o transversal a varios “módulos”, por ejemplo un proceso en el que se dan representaciones tomadas de varias modalidades sensoriales.

## DEBATES

## Cognitivistas, todavía un esfuerzo...

*Mario Borillo\**

Usted me habla. Yo lo escucho. Como usted se expresa en mi lengua materna, se refiere a cosas bastante familiares y nuestro encuentro se da en una situación que nada tiene de excepcional, le comprendo. Por lo demás, le comprendo muy bien pues a través de sus palabras adquiero cierto conocimiento que adapto a los que ya tenía...

Partiendo del flujo nervioso suscitado por las ondas sonoras en mi aparato auditivo, se me dice que fue provocada una serie de transformaciones biofísicoquímicas en los componentes elementales (¿las neuronas?) cuyo conjunto configura ciertas regiones de mi cerebro. Y se me dice que esas transformaciones han de haberse realizado según procesos de los que se sabe sobre todo que son en extremo complicados, probablemente paralelos en alto grado, para utilizar una metáfora cara a los electricistas, pero que aparentemente es la única metáfora que nos dicta nuestra imaginación frente a las imposibilidades fisiológicas, energéticas, temporales, etc. que aquejan a otras metáforas.

Si no sé gran cosa sobre la naturaleza de esas transformaciones ni sobre la manera en que se han producido, en mi vida dispongo desde hace mucho tiempo de excelentes medios para saber si su resultado global es aceptable. Por ejemplo, respondo o no respondo a sus palabras, pero la naturaleza de mis palabras y de mi conducta tratadas por los dispositivos internos de usted producen en usted reacciones, actitudes que,

\* Director de Investigación del CNRS. Director de la UA 347 lenguajes y sistemas informáticos, Tolosa.

captadas a su vez por mis ojos, mis oídos y hasta mi piel, me confirman en la sensación de que mi primer tratamiento se había realizado de manera correcta. Una tercera persona que asista a nuestra conversación puede validar o no validar esa sensación utilizando, por supuesto, sus propios medios de comprensión.

Dominique Wolton, que no teme molestar a sus amigos, me pregunta qué tiene que ver todo esto con la investigación cognitiva. Pregunta vertiginosa hasta para un científico de corte antiguo que tiene muchos principios y poca imaginación. Sería presuntuoso querer recordar en algunos renglones lo que Putnam, Quine y algunos otros han dicho en debates académicos tan sutiles como profundos, pero en fin, parece necesario que el proyecto científico se funde en su origen sobre inventarios precisos y sistemáticos de los hechos observables, según la manera de los grandes naturalistas del siglo de la Ilustración. A medida que progresa la descripción y con ella las condiciones y los objetivos de la observación, a los hombres de ciencia les parece útil y hasta hermoso, en diversas perspectivas, condensar esas descripciones alrededor de ciertos principios rectores y de ciertas propiedades estructurales expresados en un lenguaje formal. Nada veda estudiar las propiedades matemáticas. A decir verdad esto es indispensable. Recuérdese a Schützenberger cuando estableció las estructuras algebraicas que vinculan las gramáticas generativas, los autómatas abstractos y sus producciones. Llegados a este punto, la mayor parte de los investigadores reconoce que efectivamente se ha establecido una teoría (de las máquinas abstractas y de los lenguajes formales). Y esta teoría presenta aspectos constructivos completamente nuevos por cuanto describe no sólo estructuras sino también la conducta de las “máquinas” que la producen (automáticamente, es decir, por sí mismas).

Pero ¿qué decir si ya no nos referimos a los “cerebros electrónicos” sino a los cerebros sin más ni más, no a los lenguajes de programación sino a las lenguas naturales, no a los programas sino a los procesos reales que se desarrollan en nuestro cerebro? Este es el problema que interesaba a Chomsky alrededor de 1950 y que continúa estando en el centro de la investigación cognitiva.

Considerar los primeros términos de esta enumeración como idealizaciones de los segundos significó un importante

paso hacia adelante, sobre todo en el caso de la lingüística. Pero en las demás disciplinas la descripción fundadora apenas ha comenzado. Nuestros limitados conocimientos sobre la naturaleza material y la organización del cerebro y los conocimientos mucho más escasos aún sobre los procesos dinámicos que se desarrollan en él y que modifican su estado dejan una gran parte arbitraria a los modelos actuales de arquitectura material (cualesquiera que sean esos modelos) y a la naturaleza de los algoritmos que se ejecutan en ellos. Además, como lo reclamaba con razón David Marr, para evaluar en este dominio la aceptabilidad de una teoría, la distinción modelo/algoritmo debería llevarse hasta el nivel algoritmo/instrumentación. ¿Es posible imaginar la cantidad de observaciones sobre el cerebro de que habría que disponer para decidir seriamente sobre la aceptabilidad empírica de semejantes teorías?

Estas pocas observaciones elementales en modo alguno quieren ser disuasivas. Antes bien, ante el entusiasmo cognitivo, pretenden tan sólo señalar la magnitud de las dificultades que hay que afrontar. Tanto más cuanto que esas dificultades en el caso de la observación se duplican con dificultades teóricas que no son desdeñables. No cabe la menor duda de que los conceptos de memoria asociativa y de máquina “neuronal” son interesantes, por lo menos por el hecho de dar un paso hacia la investigación de cierta conformidad potencial con la organización del cerebro, y por esa razón pueden sugerir ciertos cotejos nuevos con la investigación experimental. Pero ¿qué hay que pensar de los algoritmos que se utilizan, de las relaciones muy particulares entre logicial y material que se explotan? Todos conocen la polémica que separó a Hillis y a Poggio de Hopfield y de Tank sobre la solución del célebre problema del “viajante de comercio”: hallar una solución simplemente aceptable, como hace la máquina de Hopfield, no constituye el verdadero desafío, que consiste en encontrar la mejor solución que haya. Y para llegar a ella hay que resolver problemas matemáticos reales, de modo que una máquina, para poder encontrar la solución, debe poner por obra algoritmos (probablemente paralelos) que realicen las especificaciones del modelo.

Este ejemplo suscita algunas preguntas sobre la necesidad de las teorías formales, sobre su papel, sobre su naturaleza. ¿Pueden los planos de cableado de todos los sistemas de ida y de venida reemplazar el álgebra de Boole? ¿Por cuáles univer-

sales prototípicos (subsimbólicos), si alguien es capaz de definirlos de manera estable, se pueden reemplazar los principios fundamentales de la lógica? ¿No haría falta más materia que toda la que existe en el sistema solar para que podamos esperar construir (sin otras constricciones formales que las de las máquinas “neuronales” actuales) un dispositivo que tenga la potencia y la universalidad computacionales de nuestro sistema de cálculo? Parece difícil imaginar que una modificación significativa, si no radical, de la arquitectura material básica pueda no exigir una teoría correspondiente que garantice la validez de los cálculos que puedan ejecutarse. Si esto no es así, si no fuera necesario y además difícil, mis colegas informáticos que trabajan en máquinas y en los programas paralelos “clásicos” sin duda ya habrían realizado sus proyectos de investigación o los habrían abandonado. Estamos lejos de la meta. Para volver a considerar las arquitecturas “neuronales”, ¿no es cierto que sus maravillosas cualidades de flexibilidad en el aprendizaje disminuyen cuando el número de capas de procesadores elementales aumenta, siendo así que este aumento permitiría, según parece, reducir las constricciones formales necesarias para que el proceso alcance éxito? Cruel dilema.

Ningún investigador interesado en la problemática cognitiva negará, según creo, el interés que tienen los esquemas conexionistas y algunos de los conceptos relacionados con ellos. El problema procede de la parte reservada en este enfoque al aspecto formal, del cual algunos hasta pretenden que es prácticamente inútil, pues los problemas (de discriminación, de aprendizaje) sometidos a los sistemas encuentran su solución en virtud de sus solas arquitecturas (o casi). Es lícito tener una posición muy diferente y pensar que, por algunas de las razones que antes expuse, el dominio teórico de esos sistemas y cálculos que se hagan en ellos exigirá recursos matemáticos y lógicos importantes, como lo indica la historia de la informática teórica. Una vez postulado esto, puede uno desinteresarse por este aspecto de las cosas y lograr, según se dice, algunas hermosas aplicaciones. Muy bien, pero ¿hasta dónde?

Inmensas zonas vírgenes en el mapa funcional del cerebro, marcos teóricos que tal vez haya aún que descubrir (si no creemos en la conexión de las memorias asociativas como un milagro y si no estamos seguros de que la actividad mental se reduzca a cómputos de los cuales daría cuenta la máquina de

Turing), cuestiones filosóficas ciertamente, cuando me defino en el acto de hablarle a usted, de responderle o no y cuando un tercero nos observa para decir lo que pasa... ; así quedan sugeridas algunas tareas o interrogaciones indispensables. Son monumentales. Tal vez sea sensato no afrontarlas directamente en el trabajo cotidiano, no hablar de ellas casi nunca aunque pensando continuamente en ellas. Si el rasgo genial tarda en llegar, siempre quedará la acumulación de pequeñas cosas bien hechas.

## Observaciones sobre las ciencias cognitivas

Jean-Pierre Desclès\*

Las respuestas a los cuestionarios enviados a los grupos interesados, dentro del marco de la Acción de Investigaciones Integradas "Communication" del CNRS, y las discusiones que se desarrollaron alrededor de las ciencias cognitivas me han inspirado algunas observaciones que ahora presento en la forma de trece lemas, expresados sin matices y sin argumentación.

*Lema 1.* Actualmente las ciencias cognitivas presentan más un programa concertado de investigaciones y de métodos que un conjunto homogéneo de resultados espectaculares.

Las ciencias cognitivas deben desarrollar, pues, una actividad científica partiendo de nuevos problemas bien discernidos y no defender una doctrina llamada cognitiva.

*Lema 2.* Ciertas especulaciones y ciertas analogías presurosas (como "Las relaciones entre espíritu-mind- y cerebro-brain- son de la misma naturaleza que las relaciones entre programa y ordenador"; "el espíritu es un conjunto de *logiciels* —software— del cual el cerebro constituye el hardware"; "el ordenador es un modelo del cerebro"; "el cerebro funciona como una máquina de Turing"; "cerebros y ordenadores están en una misma relación de funcionalidad"... ) son cómodas pero están demasiado cargadas de temibles trampas para alimentar un

\* Universidad de París-Sorbona, Director del Centro de Análisis y de Matemáticas Sociales, Unidad Mixta CNRS/EHESS.

verdadero trabajo de laboratorio:

Esos trabajos de laboratorio suponen reflexiones desarrolladas según los cánones de la ciencia, especialmente técnicas exploratorias sometidas a la experimentación y a la observación. Todo esto exige prudencia, lentitud, modestia y cierto ascetismo para llegar a formulaciones lo más adecuadas que sea posible a sus objetos, todo lo cual conduce a evitar falsas metáforas, con demasiada frecuencia vinculadas con un estado de desarrollo tecnológico de una sociedad.

*Lema 3.* No reducir las ciencias cognitivas solamente a los discursos pronunciados por los epistemólogos y por los filósofos sobre "la atención", "la intención", "la conciencia", sin que ello signifique excluirlos, pues la filosofía y la epistemología continúan siendo dimensiones indispensables en el estudio de la cognición.

Hasta existe un nuevo sector de la filosofía llamado "filosofía del espíritu" que se desprende de los trabajos de las ciencias cognitivas. Sin embargo, esas reflexiones se sitúan dentro de series históricas hechas de reiteradas "disputas", en las cuales intervinieron algunos "grandes espíritus", como Locke, Descartes, Spinoza, sin hablar de Aristóteles y los pensadores medievales.

*Lema 4.* Las "ciencias cognitivas" no son "ciencias nuevas", con vocación enciclopédica, con nuevas reglas normativas impuestas al quehacer científico.

Las ciencias cognitivas proponen una organización más eficaz de las tareas, una nueva configuración del saber y una arquitectura de los conocimientos ( en los que hay que distinguir conocimientos y metaconocimientos); así invitan a un enfoque interdisciplinario de problemas precisos, sin pretender sin embargo dar respuestas universales a todas las grandes cuestiones filosóficas referentes al hombre, a su inteligencia, a su lenguaje y a sus modos de adquirir nuevos conocimientos.

*Lema 5.* Las ciencias cognitivas no se reducen ni a la resolución de problemas por planificación ni a los estudios de implantaciones en redes neuronales.

La inteligencia artificial es uno de los componentes de las ciencias cognitivas. Este dominio de la informática suministra medios de programar conductas humanas que atestiguan una inteligencia cierta (comprender discursos, razonar, aprender, planificar una acción para llegar a un fin, percibir formas, representar mediante juegos de símbolos, clasificar y abstraer, producir discursos, memorizar...). Es posible, pues, estudiar la cognición de las máquinas y luego compararla con la cognición de los animales y con la cognición de los seres humanos.

Por su parte, las neurociencias participan también en el desarrollo de la cognición al precisar los modos de implantación en el seno de arquitecturas cognitivas, implantaciones de las representaciones simbólicas y de los procesos de tratamiento en los tejidos biológicos y neuronales. Además de la psicología cognitiva, la lingüística, la lógica y la informática fundamental deberían también contribuir otras disciplinas, como la descripción de las lenguas llamadas "exóticas", la etnología y la antropología mediante las cuestiones que ellas plantean. Las ciencias cognitivas no constituyen pues un conjunto cerrado de disciplinas, pues deben permanecer abiertas a los conceptos capaces de contribuir a resolver los problemas que las ciencias cognitivas formulan.

*Lema 6.* En las ciencias cognitivas, la formulación no se reduce únicamente a la construcción de sistemas formales deductivos. En las ciencias cognitivas, la simulación no se reduce a una reproducción (programada en ordenadores) de los fenómenos de conducta estudiados.

La formulación es también una matematización y una informatización de los conceptos, para definirlos mejor, para discernirlos mejor, para determinar mejor sus extensiones y sobre todo para darles una forma operativa unívoca más transmisible.

La simulación es también (y quizá sobre todo) una acción experimental que utiliza el ordenador como un órgano destinado a ejecutar las operaciones y los procesos a los que se desea dar un modelo.

*Lema 7.* Las ciencias cognitivas deben evitar todo triunfalismo, fundado mucho más en promesas difíciles de cumplir que en resultados atestiguados y en aplicaciones reales.

Se ha condenado muy justamente ese triunfalismo prematuro de la inteligencia artificial que no había logrado cumplir sus propias promesas. Pero no es menos cierto que la inteligencia artificial obtuvo resultados indiscutibles (programación avanzada, lenguajes de programación más promisorios como los lenguajes declarativos, sistemas expertos, estructuras de datos complejos como las representaciones de los conocimientos...) que ahora encuentran realizaciones industriales.

*Lema 8.* Evitar las hipótesis que sean prácticamente inverificables, ya directamente, ya indirectamente.

No tomar una hipótesis que exige verificación por un resultado adquirido que tendría fuerza de ley normativa. No confundir principios e hipótesis, pues el principio está dado por necesidad y sólo las consecuencias que se extraigan de él están sometidas a verificación. Recordar que las ciencias cognitivas tienen una base empírica, lo cual implica una teoría de aproximaciones y de errores para ajustar representaciones formales a los hechos observables y para que se haga el necesario retorno de lo teórico a los fenómenos accesibles a la observación experimental.

*Lema 9.* En un programa de formación (por ejemplo, en una formación doctoral que introduzca en el terreno de la cognición), evitar la facilidad de las lecciones panorámicas y únicamente descriptivas.

Antes bien, hay que intentar dibujar un mapa de los principales conceptos, mostrar las articulaciones conceptuales, insistir en la posibilidad de que los conceptos se deforman cuando cambian de dominio, enseñar las principales técnicas básicas (experimentales y de representación), aprender a identificar "verdaderos problemas" derivados de la cognición...

*Lema 10.* Un buen especialista de las ciencias cognitivas no deberá convertirse en "un especialista de nada sobre todo".

Deberá evitar conocer únicamente la jerga de las ciencias cognitivas sin conocer la red de los conceptos, sin saber aplicar técnicas precisas de indagación y de experimentación, sin ser capaz de recurrir a las representaciones formales adecuadas al estudio de un problema preciso. El experto en ciencias cognitivas es aquel que domina las articulaciones interdisciplinarias necesarias para resolver por completo un problema; se parece más a un arquitecto que a un técnico especializado.

*Lema 11.* El carácter interdisciplinario no debe reducirse a un discurso global y heterogéneo que así resultaría incomprendible para todos.

Sería conveniente que la comunidad reflexionara sobre reglas que guíen los discursos interdisciplinarios y sus evaluaciones.

*Lema 12.* Recordar que uno no es realmente competente y productivo en más de dos disciplinas a la vez (con raras y notorias excepciones).

Todo programa de formación debe pues tener en cuenta esta realidad.

*Lema 13.* Sería peligroso encerrarse en problemáticas de escuelas y en modas dominantes (que como todas las modas habrán de ser superadas por otras modas).

De esta manera se correría el riesgo de enceguerse y rechazar a priori nuevos enfoques, que sin embargo podrían llegar a ser más fecundos en el futuro. Las referencias internacionales son excelentes e indispensables, pero tal vez no sean ellas el único criterio para orientar y determinar los programas de investigación. Hacerse únicamente eco de las investigaciones desarrolladas “en otro lugar” no constituye un programa de investigaciones original que tenga como legítima ambición aumentar nuestros conocimientos sobre la cognición.

De la colección de **PAPELES JPG**  
en <http://padron.entretemas.com.ve>

## El informe sobre las ciencias cognitivas

*Michel Imbert \**

El informe realizado por el comité de orientación del Programa de Investigación sobre las Ciencias de la Comunicación tiene el gran mérito de existir y de ofrecer cierto número de observaciones útiles, por más que algunas sean a veces redundantes o no siempre suficientemente claras. El principal reproche que yo le haría es de otra índole y no se refiere a aspectos de redacción (por ejemplo, la utilización del término *neuropsicología*, que denota una disciplina institucional tradicional, para designar en realidad las neurociencias cognitivas). Estos aspectos son en realidad perfectibles. La principal crítica que se puede hacer a este informe es que da una imagen algo dispersa de las ciencias cognitivas. En el informe dichas ciencias aparecen como un agregado de disciplinas dispares, cada una de las cuales permanece ajena a las otras. Esta imagen refleja sin duda la situación institucional actual en nuestro país, pero sin embargo no debe hacer olvidar que las ciencias cognitivas se definen mejor por las cuestiones que plantean que por las técnicas y los instrumentos que utilizan. Esas cuestiones atañen menos a las estructuras detalladas que a las funciones globales que son capaces de realizar el cerebro y el espíritu: ¿qué es lo que nos permite ver, oír, sentir, reconocer, aprender, recordar, elegir, decidir hacer una cosa o no hacerla, en una palabra, ser inteligentes?

Desde luego, por razones prácticas, la mención de los equipos que trabajan en un campo de investigación tan vasto no puede dejar de dar la impresión de una enumeración: las

\* Profesor del Departamento de las Neurociencias de la Visión. IDN. Université Pierre et Marie Curie, París.

neurociencias, la psicología, la lógica, la lingüística y también la inteligencia artificial. Estas son en realidad subdisciplinas especializadas, a veces bien individualizadas, que en el seno de estos grandes terrenos de investigación participan de las ciencias cognitivas. Esto es particularmente cierto en el caso de las neurociencias, una gran parte de las cuales sin embargo permanece ajena a las cuestiones de naturaleza cognitiva, ya porque se interesa en funciones que no son fácilmente "calculables" (las emociones, por ejemplo), ya porque se limita al análisis fisicoquímico de los materiales constitutivos del sistema nervioso. Se podría hacer una observación del mismo tipo en los casos de la psicología, de la lingüística o de la inteligencia artificial.

En verdad, la principal tarea que nos incumbe es perforar y traspasar las fronteras que nos encierran en un coto de investigación para contribuir así a la creación de una comunidad integrada de investigadores cognitivos que compartan la misma finalidad: comprender la inteligencia. Desgraciadamente en muchas universidades europeas (y no solamente francesas) las investigaciones realizadas en ciencias cognitivas permanecen muy aisladas y son impermeables a las influencias exteriores, sobre todo las de la industria y las de la defensa, lo cual pone en general obstáculos a la constitución de esa comunidad de estudiosos.

En este momento, iniciativas nacionales como el club técnico "Ciencias Cognitivas" de la ANRT (Asociación Nacional de Investigaciones Técnicas) en el que se encuentran universitarios e industriales, o iniciativas internacionales, como el programa de las comunidades europeas BRAIN y como la "red europea de las ciencias cognitivas" del programa FAST cuya acción ha permitido incorporar "Inteligencia Artificial y Ciencias Cognitivas" en el nuevo programa ESPRIT, son otros tantos pasos dados para construir ese "espacio" y facilitar a la vez la colaboración interdisciplinaria e internacional.

## Las dos dimensiones de la cognición

*Jean François Richard\**

Las investigaciones sobre la cognición me parecen desarrollarse según dos ejes:

- Tratamiento de la información sensorial, programación y control del movimiento.
- Comunicación y tratamiento de la información simbólica.

El primer eje corresponde a cuestiones tales como:

los mecanismos de codificación y de interpretación de la información sensorial, las relaciones entre tratamientos ascendentes y tratamientos descendentes que intervienen en los procesos de reconocimiento;

las representaciones espaciales y temporoespaciales, la programación y el control del movimiento.

El segundo eje debe encararse en dos aspectos: el aspecto de la comunicación, que atañe a los intercambios de informaciones con miras descriptivas, persuasivas o de relación y el aspecto de la representación, que atañe a la estructura de los conocimientos y a sus condiciones de acceso, la manera en que los conocimientos se adquieren y se modifican, la manera en que son tratados para producir inferencias, construir interpretaciones de textos o de situaciones, definir planes de acción, etc.

Se pueden distinguir dos polos en el tratamiento de la información simbólica:

\* Profesor de la Universidad de París VIII, UER de Psicología, St. Denis.



— un polo Comunicación que se refiere a la comunicación hombre / hombre y a la comunicación hombre / máquina.

Los problemas tratados dentro del marco de la comunicación hombre/hombre son: el análisis de las comunicaciones y de las interacciones en las situaciones de diálogo, los actos de comunicación, la pragmática de la comunicación, los procesos de argumentación y las estrategias de persuasión, los mecanismos de construcción de imágenes, de elaboración de opiniones.

Los problemas derivados de la comunicación hombre / máquina se refieren principalmente a la definición de un modelo del operador y del utilizador para la concepción de relaciones de comunicación, la definición de un modelo del que aprende en los sistemas de ayuda en el aprendizaje.

— un polo Representación y Razonamiento.

Las cuestiones que se presentan aquí son:

- las formas de conocimientos y los formalismos de representación de los conocimientos, la adquisición de conocimientos mediante textos o mediante textos e imágenes, el aprendizaje por la acción y la resolución de problemas,

- la construcción de interpretaciones de textos, de consignas, de enunciaciones de problemas, de preguntas o demandas de información...

- los mecanismos de inferencia y de razonamiento que intervienen en diferentes contextos: deducción, investigación de hipótesis y diagnóstico, comprensión, planificación de acción, toma de decisión y resolución de problemas,

- el control de la actividad: evaluación de los resultados de la acción, reflexión sobre la actividad para modificar las estrategias de razonamiento y de busca de solución o también la representación de la tarea (metaconocimientos y metacognición).

Dos objetivos parecen tener prioridad en el dominio de la comunicación y del tratamiento de la información simbólica. Un primer objetivo es desarrollar formalismos que permitan representar conocimientos y hacer cálculos sobre esos conocimientos con miras a producir inferencias, planes de acción y a modificar esos conocimientos partiendo de nuevas informaciones transmitidas por enunciación o captadas en una situación que puede ser interactiva o no serlo.

Un segundo objetivo, no menos importante, es desarrollar modelos de los procesos mentales expresables en lenguajes calculables a fin de permitir las simulaciones y confrontar los resultados de la simulación con las conductas observadas.

Esto supone una observación profunda de los sujetos en situación de laboratorio o en el terreno y un análisis fino de los protocolos de observaciones recogidas a fin de identificar en el nivel individual los conocimientos utilizados y los procesos de tratamiento que entraron en juego.

Construir modelos de los procesos mentales en situaciones específicas permite precisar cómo funcionan diferentes procesos que entran en una tarea (buscar información, producir inferencias, poner en tela de juicio interpretaciones, decidir acciones o reexaminar situaciones) atendiendo a las particularidades de la situación (y a la representación que se haga el sujeto de ella) y atendiendo a las coacciones e intereses del sujeto y a lo que se exige de éste, así como a las consecuencias de sus decisiones.

Los dos ejes que hemos distinguido, tratamiento perceptivo y control del movimiento, por un lado, y representación y razonamiento, por otro, tienden a constituir líneas divisorias entre campos disciplinarios.

El primer eje es el dominio privilegiado de las neurociencias y de la parte de la inteligencia artificial que se ocupa del reconocimiento de las formas y de la robótica. El segundo eje es el dominio de la lingüística, de la pragmática y de la parte de la inteligencia artificial que se ocupa de la representación de los conocimientos y de los modelos simbólicos.

La psicología cognitiva está recorrida por esa línea divisoria que se manifiesta en el debate sobre la modularidad. Los partidarios de la modularidad se encuentran en el primer eje. Conciben los módulos como unidades de tratamiento independientes las unas de las otras y la respuesta depende de informaciones específicas, sin tener en cuenta otras informaciones que existen en el sistema cognitivo y que podrían ser pertinentes, pero a las que el módulo no tiene acceso.

Fodor es el teórico de una concepción modular sobre el funcionamiento cognitivo. Los criterios que da Fodor de la modularidad son: el carácter específico, el carácter irreprimible, la impermeabilidad a los procesos centrales, la rapidez, el encapsulamiento de la información, es decir, el hecho de que no haya acceso a las representaciones de alto nivel pues se trata de una arquitectura nerviosa fija de disfuncionamientos específicos y de un desarrollo ontogenético.

Si seguimos a Fodor, hay dos formas de funcionamiento cognitivo:

- los módulos, caracterizados por el encapsulamiento de la información, en los que Fodor sitúa la percepción, la comprensión;
- los procesos centrales, que no están encapsulados y de los que procede la inferencia.

La diferencia de ambas formas es muy grande pues, según Fodor, la ciencia cognitiva se detiene en los módulos: no es posible construir teorías que simulen los procesos centrales porque éstos no son locales ni pueden por eso aislarse mediante el análisis experimental.

Parece difícil de negar la existencia de tratamientos modulares. El problema está en saber cómo éstos se sitúan dentro del funcionamiento cognitivo y qué parte tienen en él. Los tratamientos modulares que se han puesto de manifiesto atañen a tratamientos ascendentes que se sitúan en un nivel infrasemántico e infradecisional. Tienen que ver con operaciones elementales cuyos productos están muy lejos de los resultados de la actividad cognitiva, a la que se llama de comprensión o de resolución de problemas.

Consideremos el caso del lenguaje. Los dominios en que el enfoque modular tuvo mayor éxito son la segmentación del lenguaje hablado y escrito (letra o sílaba) y el acceso al léxico. En estos dominios se han manifestado efectos de independencia en relación con el contexto semántico, efectos que son los que uno espera de tratamientos de tipo modular. Por ejemplo, se ha mostrado que en el caso de palabras ambiguas eran activadas varias significaciones y que sólo al cabo de algunos centenares de milisegundos se había seleccionado la interpretación apropiada al contexto.

Hay que observar que el acceso al léxico se sitúa en un nivel infrasemántico si se admite (lo cual actualmente ya no se discute) que el nivel de base de la significación es la predicación, es decir, la estructura predicado-argumento, que no siempre corresponde a la proposición gramatical. Es allí donde está la unidad de base, en el nivel más elemental en el que se puedan atribuir las calificaciones de verdadero, falso, verosímil.

Los procesos cuyo estudio deriva del segundo eje de investigación utilizan los resultados de tratamientos modulares, eso es evidente, pero asimismo hacen entrar en juego cantidades

considerables de inferencias que se basan o bien en conocimientos almacenados en la memoria o bien en razonamientos de naturaleza inductiva o deductiva. Las investigaciones sobre la resolución de problemas son instructivas en este aspecto. Para explicar la índole de los ensayos de solución que se intentan hay que considerar la interpretación que el sujeto tiene de la situación. Para comprender cómo se construye esa interpretación hay que explicar cómo se realizan las atribuciones elementales de significación. Actualmente la idea que se destaca es la de que esas atribuciones apelan a numerosos conocimientos de naturaleza muy diversa. Tenemos buenas razones para pensar que la atribución de las significaciones que podemos llamar léxicas depende en realidad de los conocimientos de procedimiento de que dispone el sujeto tocante a la tarea a la que está entregado.

Existen diferencias importantes en las investigaciones que se desarrollan según cada uno de estos ejes, diferencias desde el punto de vista de las problemáticas, de los modelos y de los métodos.

Lo cierto es que la división que tiende a establecerse es de naturaleza más ideológica que heurística.

Lo discutible no es la existencia de tratamientos que tienen los caracteres de la modularidad; es una concepción de la modularidad derivada de las tesis de Fodor, quien asimila lo cognitivo a los tratamientos modulares y esto constituye, por cierto, un obstáculo al análisis de los procesos de tratamiento que atañen a los saberes y a las significaciones. Numerosas cuestiones importantes se sitúan en la intersección de estos dos ejes. Tales cuestiones son por ejemplo: las interacciones entre los tratamientos de tipo ascendente y los tratamientos de tipo descendente, el tratamiento de las informaciones simbólicas espacializadas (esquema, imagen), las relaciones entre programación de la acción y programación del movimiento, los mecanismos de activación de los conocimientos, las relaciones entre modelos simbólicos y modelos conexionistas. Es fundamental no pasar por alto ninguna de estas dos dimensiones de la cognición.

## ¿Qué es lo cognitivo?

Jean-Paul Codol\*

Por supuesto, es imposible condensar en tan poco espacio una concepción de lo cognitivo que no parezca demasiado truncada. Es imposible hacer justicia a la complejidad de los problemas, a la diversidad de los puntos de vista y a sus relaciones, a las cuales soy particularmente sensible por las funciones que desempeño en Aix, en el Centro de Investigaciones de Psicología Cognitiva.

Pero deliberadamente he decidido, en lo que respecta a esta cuestión, que es posible trazar líneas muy amplias de un marco de nociones y señalar algunas grandes categorías de temas de investigación relacionadas con ellas.

Evidentemente, tengo conciencia de la excesiva generalidad de todo lo que sigue. Sin embargo he intentado indicar las perspectivas en que se sitúa una buena parte de los trabajos realizados en el dominio que es el mío —el de la psicología social experimental— y que tienen relación con el estudio de la cognición.

1. Comenzaré aludiendo a algo de historia y mencionando a tres de los padres fundadores de la psicología social moderna: Lewin, Heider y Asch. Ellos me permitirán indicar ya perspectivas, temas de investigación —y por lo tanto una concepción de lo cognitivo— y recordar también que, de todas las subdisciplinas de la psicología, la psicología social es sin duda aquella cuyo interés por los fenómenos, las estructuras y los procesos cognitivos es el más antiguo y cuya tradición en este terreno es la mejor anclada.

\* Director del Centro de Investigaciones de Psicología Cognitiva, unidad asociada al CNRS, Aix-en-Provence.

Kurt Lewin fue, pues, en los comienzos de la década de 1930 —hace unos sesenta años— uno de los primeros en afirmar que la conducta individual o social no depende únicamente del ambiente material, de las situaciones o de los estímulos, sino que además depende mucho de la manera en que los individuos integran en su sistema cognitivo informaciones sobre ese ambiente, sobre esas situaciones o sobre esos estímulos y de la significación que estos factores tengan para los sujetos.

Por su parte, Heider tomó de las teorías de la *Gestalt* la idea de que las unidades perceptivas y cognitivas son fuentes de un sentimiento positivo. De manera que cuando se presentan incoherencias o contradicciones, los individuos se esfuerzan por reducirlas. Observemos que en la perspectiva de Heider esas incoherencias o contradicciones no se reducen por obra sólo de la actividad mental de los individuos. Muy por el contrario, Heider las considera como *fuerzas de motivación* que desencadenan y orientan la conducta y el comportamiento. Con esta idea se pueden relacionar las numerosas *teorías de la coherencia*.

Heider, que creía en el valor heurístico de las *psicologías ingenuas* (llamadas del sentido común) para construir una psicología científica es también el iniciador de trabajos de investigación sobre los *sistemas de pensamiento* y sobre los *esquemas cognitivos y sociocognitivos* que los subtienden y los organizan. ¿Cómo explican los individuos los sucesos de que son actores u observadores? ¿Cómo, partiendo de las informaciones de que disponen, infieren otras informaciones sobre lo que no saben? Heider abre así el camino de innumerables trabajos sobre los *procesos de inferencia* —y muy particularmente sobre la inferencia causal— y las *teorías de la atribución*.

Por fin, le debemos a Asch dos tradiciones distintas de investigación, una sobre la influencia social y la otra sobre la percepción de las personas.

En lo referente a los procesos de influencia, Asch los concebía esencialmente atendiendo a un *conflicto cognitivo*, conflicto entre informaciones de las que dispone un sujeto por experiencia personal (informaciones perceptivas directas, por ejemplo) y las informaciones que le transmite su ambiente social.

En cuanto a la percepción de las personas, Asch es el iniciador de toda una línea de trabajos que sobrepasan amplia-

mente este terreno y que se refieren mucho más generalmente a los *procesos de integración de la información* y a la manera en que los sistemas psíquicos se forman *impresiones* sobre objetos. Cuestiones como las de la primacía de la información, de su carácter central en una representación de conjunto, de su complejidad, de su diferenciación, etc. constituyen otros tantos temas muy copiosamente estudiados que tienen relación con los campos de investigación iniciados por Asch.

Las ideas de Heider y de Asch contribuyeron asimismo mucho al nacimiento y al desarrollo de trabajos sobre las *teorías implícitas* y especialmente, tratándose de la percepción de las personas, contribuyeron a los trabajos sobre lo que se ha llamado *teorías implícitas de la personalidad*. Aquí se trata de discernir los procesos en virtud de los cuales se elaboran las representaciones implícitas que se forman los sujetos de una determinada persona. Pero también aquí el concepto de teoría implícita sobrepasa ampliamente el tema de la percepción de las personas y hasta de los objetos sociales; en realidad, dicha teoría atañe a todo estudio de la representación.

Como se ve —para no tomar más que estos pocos ejemplos—, hay pocos dominios de la psicología social actual que no tengan una orientación cognitivista de larga data (más de cincuenta años para algunos de nosotros). Al comienzo numerosos trabajos de psicología cognitiva se desarrollaron primero en el seno de la psicología social, siendo así que no presentaban ningún carácter particular que los relacionara de alguna manera natural con esta subdisciplina. Por eso legítimamente se ha podido sentir asombro de que la psicología social asimilara temas de investigación como los que se refieren, por ejemplo, a la disonancia cognitiva, a la formación de las impresiones, a las teorías implícitas, a las inferencias, a los esquemas y sesgos cognitivos o más recientemente a la teoría de los *scripts*, etc. La mayor parte de estas teorías pretenden, en efecto, describir procesos cognitivos intraindividuales ignorando su contenido y su origen social. Y esta asimilación, a decir verdad, parece más el resultado de avatares históricos producidos en el desarrollo de la investigación que de una necesidad teórica fundamental. Sin duda hay que ver aquí precisamente la indicación del poco interés que mostraba la psicología general de entonces por lo cognitivo, que abandonaba por entero a la psicología social. En todo caso, la psicología social participó mucho más de lo que en

general se cree en lo que se ha dado en llamar la “revolución cognitiva” en psicología.

Con estas pocas palabras de historia, ya he mencionado algunos grandes temas de investigación de una psicología social que se interesa en la cognición: representaciones subjetivas internas, conflictos cognitivos, incertidumbre, incoherencia y su relación con el desencadenamiento y la orientación de la conducta, trabajos sobre los sistemas de pensamiento y sobre los esquemas cognitivos y sociocognitivos que están en su base y los organizan, procesos de inferencia y de atribución, procesos de influencia, formación de las impresiones y procesos de integración de la información, teorías implícitas, etc.

2. Para presentar las cuestiones de otra manera, partamos ahora de una consideración trivial y de una definición.

La trivialidad es la siguiente: las informaciones a que nosotros estamos continuamente sometidos son en extremo numerosas y de naturaleza muy variada; algunas nos llegan directamente de nuestros sentidos, otras de nuestra memoria, otras, transmitidas o no por el lenguaje, nos son suministradas por obra de las relaciones que mantenemos con los demás, individuos o grupos.

La definición es la siguiente: la *cognición* designa el conjunto de actividades por las que todas esas informaciones son tratadas por un aparato psíquico, cómo éste las recibe y las selecciona, cómo las transforma y las organiza, cómo construye así *representaciones* de la realidad y elabora conocimientos.

Creo que en esta simple definición tenemos lo esencial del programa de trabajo de los psicólogos que se interesan por la cognición. Y en cada una de las etapas del tratamiento de esta información multiforme (en la recepción, en la selección, en la transformación, en la organización, en la construcción de representaciones y en la elaboración de conocimientos), podemos discernir toda una serie de problemas, de temas de investigación, de teorías elaboradas, de trabajos efectuados o en curso.

Agreguemos —y esto es importante— que este estudio del tratamiento de la información por un aparato psíquico puede realizarse en diversos aspectos que incluyen por ejemplo la génesis, el desarrollo y el aprendizaje individual y colectivo de las actividades y de los procesos cognitivos, de su activación en situaciones particulares en los contextos en que se dan, etc.

En verdad, me parece que esto es lo esencial y creo que podría quedarme en este punto. Sin embargo, por supuesto, podemos ir un poco más lejos y precisar algo más. Aquí no puedo hacerlo más que en una sola dirección que me parece particularmente importante en psicología social.

De la información tratada resultan, en efecto, *saberes*. Estos nos permiten comprender nuestro ambiente, adaptarnos a él y obrar sobre él. La cognición tiene pues esencialmente una función de adaptación y de regulación. Identificar y reconocer los múltiples *objetos* del ambiente, darles un valor y un sentido son actividades fundamentales que están en el centro de todo proceso cognitivo.

En esta perspectiva, dos ideas por lo demás ligadas entre sí me parecen estar en el origen de nuestras concepciones sobre la cognición.

Según la primera, para un aparato psíquico dado un objeto sólo adquiere sentido, valor, significación por *comparación* con otros objetos ya conocidos, identificados y significativos. En general, se considera que esta correspondencia simplificadora pero portadora de sentido entre lo nuevo y lo ya adquirido se asegura en el curso de *un proceso de categorización*, en un sistema clasificatorio que tiene la función de seleccionar y de identificar la información recibida.

Todas las cuestiones vinculadas con procesos de categorización, con la elaboración de diferentes tipos de categorías, con el lugar que ocupan objetos dentro de esas categorías, con las relaciones intracategorías e intercategorías, con la organización de las categorías en un sistema global, con operaciones efectuadas por el aparato psíquico dentro de ese sistema (y que le permiten especialmente construir por abstracción categorías o nuevos objetos que sólo pueden tener una remota correspondencia con la realidad material de objetos concretos), etc., todas estas cuestiones, pues, son importantísimas para quien se interese por la cognición y por la elaboración de los saberes.

Importa declarar aquí que las categorías no son nunca *estrictamente definitorias*. Partiendo de la experiencia personal o sufriendo la acción de normas sociales, las características y propiedades de los objetos están siempre *valorizadas* de cierta manera por cada individuo y están así vinculadas con *esquemas de conducta* que acercan al objeto o lo evitan. Las relaciones

entre cognición y conducta individual (ya se trate de conducta motora, de conducta lingüística, de la expresión de actitudes o de opiniones, etc.) así como las relaciones entre cognición y conducta social constituyen uno de los temas particularmente importantes de los trabajos sobre lo cognitivo.

La segunda idea es la de que la categorización y la significación sólo son posibles sin embargo porque vivimos en un ambiente físico y social relativamente estable. Porque los mismos objetos poseen características subyacentes más o menos invariables es posible que un sistema psíquico los identifique. Esta *constancia de los objetos* nos remite a la *coherencia del sistema de categorización* de los individuos, que puede manifestarse de diversos modos. La constancia de las formas es un ejemplo bien conocido. Otro ejemplo: no aceptamos de un objeto ya identificado características inesperadas o contradictorias en comparación con las características que están almacenadas en nuestra memoria.

Precisamente esta estabilidad del ambiente y esta relativa coherencia psicológica son las que pueden permitir al individuo anticipar, hacer predicciones sobre los objetos, sobre las personas, sobre los sucesos sociales, etc. En todo los casos puede considerarse que el individuo construye así *teorías ingenuas* que lo ayudan a comprender e interpretar los objetos y los hechos, y por lo tanto, a guiar su conducta. Entre estas teorías, según ya dije, las teorías implícitas o los procesos de formación de las impresiones son particularmente importantes en psicología social. Los procesos de inferencia y de atribución derivan de esas elaboraciones ingenuas (por las cuales atribuimos un sentido al mundo que nos rodea) y también de los *procesos de generalización*.

Los procesos de generalización también se basan en la idea de una permanencia relativa del ambiente. Partiendo de experiencias pasadas, personales o sociales, relativas a objetos dados, cada uno de nosotros tiene, en efecto tendencia a generalizar, tanto en el tiempo como en el espacio, las características y propiedades de esos objetos. Y esta generalización a su vez guía los procesos de selección, de integración y de organización de la información ulterior relativa a esos objetos. Aquí han de considerarse los trabajos sobre los planes y esquemas: el *plan* es en efecto el aspecto operativo de ese proceso de generalización y el *esquema* es su aspecto figurativo. El plan es, pues, de

alguna manera la forma general de un proceso de generalización y el esquema es su contenido particular en determinadas circunstancias.

Planes cognitivos (que conducen a *sesgos* cognitivos) organizan así las representaciones que uno se hace de este o aquel aspecto del ambiente, de un objeto o de un hecho. Fundamentalmente, planes y esquemas constituyen expectativas relativas a la estructura de la información que se encuentra. Ellos guían la codificación de los estímulos así como apelan a la información almacenada en la memoria.

Citemos también los trabajos sobre los *scripts*, es decir, esos planes o esquemas particulares de programas que definen una secuencia de acción, etc.

Entre muchas otras éstas son algunas orientaciones para considerar fenómenos cognitivos. Valiéndome de algunos ejemplos quiero insistir en dos puntos:

— Primero, *lo cognitivo no se reduce únicamente a operaciones intelectuales*. El estudio de esas operaciones sólo constituye una parte que en realidad considero modesta de las operaciones cognitivas.

— Segundo, *la parte de lo afectivo y de la motivación* en el estudio de los procesos cognitivos. La afectividad, las emociones, la valorización, etc. intervienen realmente en todos los estadios de la actividad cognitiva: en la selección de las informaciones, en su transformación, en su organización, etc.

Evidentemente la cognición no puede reducirse a un conjunto de operaciones formales, a una alquimia más o menos compleja de elementos neutros realizada independientemente de la “vivencia psicosocial” de las personas. Verdad es sin embargo que tener en cuenta esas “vivencias psicosociales” plantea tremendos problemas en la investigación, especialmente cuando se trata de experimentación.

3. Quisiera concluir con dos observaciones referentes sobre todo a las relaciones entre la psicología y las otras ciencias de la cognición.

a) La primera es una observación de vocabulario. A decir verdad, la expresión “ciencias cognitivas, construida según el modo gramaticalmente aberrante de la expresión “ciencias económicas” (que sin embargo a veces son bien costosas) es un problema. Como se trata de un conjunto compuesto por las contribuciones de varias disciplinas científicas a un mismo

campo de estudio, sería mejor que las llamáramos más bien “ciencias de lo cognitivo” o también “ciencias de la cognición”. Se habla efectivamente de “ciencias del lenguaje” (y no de ciencias lingüísticas), se habla de “ciencias de la educación” (y no de ciencias educativas). Y cada vez se habla menos de ciencias políticas para decir en cambio “ciencias de lo político”. No es demasiado tarde para rectificar un empleo de la expresión que probablemente tenga más implicaciones epistemológicas de lo que parece.

b) La segunda observación se refiere al hecho de que son posibles evidentemente muchas ciencias de la cognición. Sólo difieren entre sí por el *sistema de tratamiento* de la información que consideran. Sistemas biofísicos y sistemas sociales, sistemas técnicos y sistemas formales, toda una gama de sistemas de tratamiento que conduce a la construcción de representaciones y a la elaboración de conocimientos es digna de ser estudiada, de manera que es posible toda una serie de “ciencias de la cognición”. Cuando se trata del sistema *psíquico*, situado en la intersección de lo biofísico y de lo social, lo que hay que considerar desde luego es la *psicología* en sus relaciones con las neurociencias, por un lado, y con las ciencias del hombre, por otro.

Es evidente que al interesarse específicamente por el sistema psíquico, la psicología de la cognición no puede reducirse a ninguna otra ciencia de la cognición. Es claro sin embargo que la elaboración de un campo integrado de investigación de las ciencias de la cognición constituye hoy un empeño científico y social considerable, que por lo demás se impone asimismo por consideraciones de orden tecnológico y económico y por la urgencia de realizaciones en diferentes dominios: desarrollo de sistemas informáticos llamados “inteligentes”, relación de intercambio entre el hombre y esos sistemas, cuestiones planteadas por el reconocimiento de las formas (visuales, auditivas, táctiles, etc.), la robótica, los conjuntos de imágenes, los sistemas expertos, la ingeniería didáctica, la concepción asistida por ordenador, etc.

En esta integración, las técnicas de formulación (matemáticas, lógicas, lingüísticas, informáticas, psicológicas) deben tener una parte fundamental: sólo ellas pueden permitir a la vez representar sistemas de tratamiento de la información que se sitúan en diferentes niveles semánticos y articular entre sí esos sistemas.

En todos los casos, la parte que tiene la psicología en las ciencias de la cognición es evidentemente central, tanto por los modelos de actividad mental que ella propone, como por la larga experiencia que tiene de las técnicas de validación en la experimentación humana. Me parece, pues, que los psicólogos nada tienen que temer si ven a algunos de ellos empeñarse resueltamente por este camino de una investigación integrada. (Digo bien “algunos de ellos” pues, como disciplina, la psicología sobrepasa y mucho a la psicología de la cognición.) Sin embargo la condición para alcanzar éxito en esta integración es la de que los psicólogos sepan hacer reconocer su propia competencia y el carácter específico de su enfoque. Nada me parecería más lastimoso que la constitución de un campo integrado de investigación que condujera a reducir la psicología a la condición de una simple *prestataria de servicios* en beneficio, por ejemplo, de las ciencias físicas de los ingenieros (informática, inteligencia artificial, robótica o alguna otra) que ciertamente saben mucho menos que nosotros en materia de cognición. Avancemos pues resueltamente y sin complejos.

## ¿ Qué es lo cognitivo?

Guy Tiberghien\*

Para no dar una respuesta demasiado abrupta a una cuestión que manifiestamente lo es, prefiero abordar el tema de soslayo. Trataré primero de definir brevemente las opciones programáticas de los psicólogos que aspiran a estudiar lo cognitivo. Esto equivale, por lo tanto, a definir en una primera etapa la psicología cognitiva y, en una segunda etapa, a señalar algunos de los problemas que se desprenden de dicha definición.

### 1. ¿ Qué es la psicología cognitiva?

La definición supone que no se puede edificar una ciencia psicológica partiendo solamente de la observación de las relaciones funcionales entre las situaciones varias y las modificaciones de conducta ( $R = f(S)$ ). En otras palabras, la psicología cognitiva procura superar las limitaciones del programa conductista. Para hacerlo postula lo siguiente: 1) que entre S y R se intercala un conjunto de procesos hipotéticos (módulos de tratamiento cuyos productos intervienen en la determinación causal de la conducta; 2) esos procesos aplican operaciones de transformación a informaciones de entrada y generan informaciones de salida; 3) esos procesos están organizados según una arquitectura que debe dilucidarse (secuencial estricta, paralela, secuencial paralela, en cascada...); 4) dichos procesos están organizados en niveles de tratamiento y son función de la naturaleza y de la forma de las informaciones recibidas así como de la naturaleza de las operaciones que realizan. En otros

\*Profesor del Laboratorio de Psicología Experimental (UA CNRS 0665) Universidad de Grenoble II.

términos, la psicología cognitiva se adhiere a una concepción popularizada conocida con el nombre de STI (Sistema de Tratamiento de la Información). Esto implica que los fenómenos psicológicos son el producto de un sistema, sistema ciertamente complejo pero que debe estudiarse como todos los demás sistemas (considérese la teoría de los sistemas). Los módulos de este sistema psicológico, lo mismo que los módulos de todo otro sistema, tratan entradas y salidas que pueden manifestarse en diferentes formas (movimiento, energía, información) en función del lugar que ocupan dentro de la jerarquía del sistema. La psicología cognitiva está fundamentalmente interesada en las entradas y salidas de tipo “informático”. Pero el concepto de información es, en muchos aspectos, inasible.

## 2. Información y cognición

Se puede entender este concepto en el sentido de la teoría de la información y definirlo desde el punto de vista de la incertidumbre (entropía) o asimilarlo a un “contenido”, es decir, a una semántica. En el primer sentido del término, el concepto de información nos remite a las coacciones sintácticas, de manera que se aplica a fenómenos muy diversos de cualquier naturaleza; en la segunda acepción, la definición semántica del contenido “informativo” delimita necesariamente un dominio de validez. En psicología cognitiva se hablará de representación “mental” o de estados “mentales” para designar la información específica transmitida por el sistema psicológico. Se puede entonces “reducir” la psicología cognitiva al siguiente esquema  $R = f(S \times M)$ , en que M designa esas representaciones mentales. En todo caso, esta solución me parece por lo menos parcialmente insatisfactoria. En efecto, sospecho que cuando se habla de representación mental algunos psicólogos piensan sobre todo en las informaciones de naturaleza simbólica que somos capaces de evocar conscientemente y de manipular intencionalmente.

### 2.1 El “todo cognitivo”

Ciertamente lo que en el hombre llama la atención con razón es la importancia de esos fenómenos mentales de naturaleza simbólica en los que son posibles operaciones intencionales (lenguaje, inferencia, creencia). Pero estos fenómenos de orden

superior son ellos mismos productos derivados del funcionamiento de otros módulos de tratamientos situados en un nivel inferior de la jerarquía de los tratamientos. Estos fenómenos pueden considerarse subsimbólicos y son automáticos. La designación “cognitivo” no debe limitarse pues a los primeros sino que también puede aplicarse a los segundos *en la medida en que la relación entre los dos niveles pueda especificarse*. Una ilustración elemental puede servir para aclarar este punto:

“Los modelos recientes del reconocimiento de rostros suponen que ‘el cálculo’ de la familiaridad de un rostro se efectúa por obra de un módulo de tratamiento automático cuyo funcionamiento no es modificado por los procesos de nivel superior (identificación, memoria semántica y episódica). Si esto fuera así (¡R. Bruyer, A.C. Schreiber y yo tratamos de demostrar experimentalmente lo contrario!), no sería menos legítimo admitir que el mecanismo de estimación de la familiaridad de un rostro es de naturaleza cognitiva y que su producto es indiscutiblemente una forma de conocimiento, por más que no sea una forma ni simbólica ni intencional (¡trate el lector de no experimentar un ‘sentimiento’ de familiaridad en presencia del rostro de su madre o trate de describir precisamente ese sentimiento!)”.

Algunos investigadores pueden considerar que se corre un riesgo extendiendo el atributo cognitivo a niveles de tratamiento de la información cada vez más elementales, porque en definitiva todo tratamiento de información podría considerarse como cognitivo. Sobre este particular se cita frecuentemente la información genética, de la que se declara inmediatamente que por cierto no es de naturaleza cognitiva. Es verdad, pero la cuestión sin duda no es tan sencilla. Hay que recordar que después de los espectaculares progresos realizados en la dilucidación del código genético, muchos psicofisiólogos sostuvieron la hipótesis de que la información mnésica relativa a las relaciones entre estímulo y conducta podía codificarse en la estructura molecular del ARN y de las proteínas sintetizadas por la neurona (véase Hyden, Ungar, etc. de mediados de la década de 1960). Después de los primeros resultados experimentales positivos, esta hipótesis fue rechazada. Sin embargo es claro



que si esta hipótesis hubiera sido verificada, la índole del ARN (que no es un fenómeno cognitivo) habría quedado profundamente modificada en relación con los datos referentes a la conducta y a la cognición en el dominio de la psicología de la memoria. En el mismo orden de ideas, ahora numerosos investigadores exploran otro camino con miras a poner en relación las modificaciones estructurales de la morfología sináptica (parámetros de los contactos sinápticos) con el aprendizaje y la memoria. También aquí, si quedara aclarada la “sintaxis” de estas transformaciones estructurales, no podría ya pasarse por alto este nivel de descripción en el análisis de los fenómenos psicológicos de adquisición o de memorización. Todo esto sugiere niveles diferenciados de la información y, a mi juicio, lo decisivo son las relaciones lógicas (en el sentido fuerte del término) y empíricas que somos capaces de discernir entre esos niveles.

## 2.2 Una cognición mínima

Otros investigadores adoptan a veces una línea de razonamiento que puede conducir a conclusiones radicalmente opuestas. Estoy de acuerdo con lo esencial de las tesis desarrolladas por Fodor (modularidad, encapsulamiento de ciertos procesos, distinción entre módulos de transducción, módulos periféricos y procesos centrales). Sin embargo, dos cuestiones me separan del análisis fodoriano: a) los procesos centrales engendran representaciones de alto nivel (simbólicas, conscientes, deliberadas) que resultan de la interacción entre los productos de los sistemas periféricos tabicados y los contenidos de la memoria (fijación de creencias, inferencia, “pensamiento”). Sin embargo, según Fodor (1983), los procesos centrales no pueden estudiarse científicamente y si, según él, no sabemos nada de la psicología ni de la neuropsicología del pensamiento eso se debe a que “aquí no hay nada que saber”. Semejante actitud equivale, de hecho si no ya en principio, a excluir una gran parte de los fenómenos mentales del campo de lo cognitivo y limitar esta designación únicamente a los módulos periféricos y a los transductores. Y esto me parece metodológica y epistemológicamente discutible, tanto más cuanto que consideraría más fructífero disociar, por

un lado, el sistema de inferencia (que está mucho mejor comprendido y hasta informáticamente simulado de lo que da a entender Fodor) y, por otro, el sistema de elaboración de las creencias (que todavía continúa siendo enteramente opaco y que sin duda es difícil de simular); b) si es sin duda correcto postular que los mecanismos de transducción son enteramente “compilados” y “no penetrables” por procesos cognitivos de nivel más elevado a causa probablemente de la importante función de adaptación al ambiente que dichos mecanismos aseguran, resulta mucho más discutible que todos los procesos periféricos sean compilados y no penetrables. Algunos tienen sin duda esta propiedad (percepción de relaciones en el espacio de 3D), pero la cuestión es mucho más difícil de resolver en el caso de otras facultades que se suponen “verticales” (por ejemplo, el reconocimiento de rostros). En realidad, la posición de Fodor conduzca tal vez a sobreestimar los procesos del tipo *bottom-up* y a minimizar los procesos del tipo *top-down*. Una breve ilustración puede mostrar la complejidad de la cuestión:

Conocida es la ilusión llamada de Poggendorf y se sabe que no se puede dejar de ser sensible a ella aun cuando uno se dé cuenta de que se trata de una ilusión; es pues razonable suponer que los procesos cognitivos que están en el origen de esta deformación perceptiva se encuentran ‘tabicados’ y son insensibles a los procesos centrales responsables de las inferencias y de las creencias engendradas por esta situación; sin embargo, basta con introducir el dispositivo de la ilusión en un contexto significativo para que ella se disipe como por encanto, ¡el sujeto ya no puede dejar de no ver la ilusión! Los módulos que entraban en juego no eran quizá tan autónomos como parecían. Esta dificultad experimental se ve agravada por obstáculos de orden lógico, pues demostrar que un módulo está encapsulado consiste en tratar de demostrar que un módulo de nivel superior no puede modificar su funcionamiento, lo cual equivale a querer demostrar (¡desesperadamente!) la nulidad de la hipótesis.

## 2.3 Cognición y niveles de cognición

Todo esto me conduce al problema inicial. Sugiero pues definir lo cognitivo, de manera muy general, como el conjunto de fenómenos resultante de la codificación, del almacenamiento y de la manipulación de la información por obra del sistema llamado en el mundo animal sistema nervioso central. La psicología cognitiva estudia precisamente esos procesos y sus productos en la medida en que éstos dan sentido a la conducta observada: la psicología cognitiva es en cierto modo la (psicológica) lógica de la psicología de la conducta. Los módulos de tratamiento postulados por la psicología pueden estudiarse en diferentes niveles: a) finalidad de los cálculos efectuados (la lógica y la matemática son los instrumentos que sirven para elaborar estas teorías “computacionales”); b) medios funcionales, algorítmicos y heurísticos, por ejemplo, aplicados para realizar esas finalidades (la psicología cognitiva, la neuropsicología cognitiva, la psicolingüística contribuyen a dilucidar esos medios mediante la experimentación, la simulación y la emulación); c) por fin, la neurofisiología trata de comprender y de “reconstruir” las arquitecturas que realizan *físicamente* las operaciones definidas en el nivel anterior. Sería particularmente vano defender la preeminencia de uno u otro de estos enfoques (si no es por razones institucionales poco confesables). Lo cognitivo resulta de la interacción entre estos diferentes niveles de análisis. Una metáfora puede ayudar a comprender este punto de vista: la “movilidad” de un automóvil se encuentra a la vez y de manera diferente en la chispa que se produce dentro del cilindro, en el desplazamiento vertical del pistón y en la rotación de los neumáticos. De la misma manera la “cognitividad” se encuentra a la vez y de manera diferente en la modificación bioquímica de las conexiones sinápticas, en la arquitectura neuronal y en la identificación perceptiva de un rostro familiar. Lo simbólico no existiría sin lo subsimbólico que lo genera, y las microcogniciones son tan “cognitivas” (en otro nivel) como las macrocogniciones. Por lo demás, creo que esta toma de conciencia de la necesaria relación entre estos niveles de explicación caracteriza el período actual, como lo atestigua el desarrollo de relaciones disciplinarias nuevas: neurociencias cognitivas, neuropsicología cognitiva, neurocomunicaciones, neuroinformática y neuroetología, por ejemplo.

## 3. De la psicología cognitiva a la ciencia cognitiva

De la manera en que he tratado de definir lo cognitivo (o, mejor dicho, ¡de hablar de lo cognitivo!) se desprende que el “programa” de la psicología cognitiva sólo puede cobrar todo su sentido si se reconocen sus vínculos con la lógica matemática (formulación de las relaciones de signo y símbolo, por ejemplo) y con las neurociencias. En realidad, el estudio de los fenómenos cognitivos no puede ser el objeto exclusivo de una sola disciplina; por el contrario, dicho estudio exige una acrecentada cooperación entre varias disciplinas. En este estadio me parece, pues, plenamente justificado hablar de ciencias de la cognición o de ciencias cognitivas o de ciencias del conocimiento (pues el debate terminológico resulta aquí relativamente secundario). Y no hay que olvidar que un nuevo actor entró en escena desde hace algunos años; me refiero a la inteligencia artificial. La aparición de máquinas “informacionales” capaces de realizar, con otros medios, ciertos actos de conocimiento realizados por el sistema nervioso iba a crear una situación sin precedentes. Los diferentes niveles de análisis de los fenómenos cognitivos (finalidad del cálculo, medios, realización material) pueden asimismo discernirse en los artefactos informáticos cuya tecnología somos capaces de dominar. Pero si la informática fundamental se asignaba el simple objetivo de aplicar algoritmos a datos indiferenciados, la inteligencia artificial mostrará la eficacia de procedimientos de inferencia (heurísticos y programables) aplicados a datos estructurados (¿conocimientos?).

### 3.1 El lugar de la inteligencia artificial en las ciencias cognitivas

Dos observaciones sobre la inteligencia artificial: a) se trata, según la hemos definido, de una auténtica ciencia experimental; b) puede aportar a la psicología cognitiva diversos formalismos capaces de facilitar la unificación teórica de datos experimentales (papel de las redes semánticas en la descripción de los conocimientos declarativos, por ejemplo). Nuevas representaciones de la información pueden asimismo engendrar nuevas arquitecturas que apunten a un relativo neuromimetismo

o psicometismo. Así lo atestigua, por ejemplo, el importante desarrollo de modelos conexionistas y neoconexionistas en el dominio de las neurociencias (modelo neuromimético del cerebelo, del hipocampo, modelos neuroetológicos de las conductas y de las estrategias de captura; c) la inteligencia artificial ofrece sistemas artificiales de manipular la información cuya variedad arquitectónica va en aumento (ordenadores von Neumann, hipercubos, “máquinas de conexión”...). Estos sistemas permiten una emulación \* cada vez más precisa de modelos propiamente psicológicos cuyo interés heurístico está reconocido.

A manera de ilustración, estamos actualmente empeñados (Schreiber, Robert, Rousset, Tiberghien y Wang) en un intento de emulación en VAX de un modelo psicológico de reconocimiento de rostros —HYPERFACE 1— cuya instrumentación en lenguaje C obliga a precisar numerosos mecanismos responsables de este o aquel efecto observado experimentalmente, a delimitar mejor los factores críticos que entran en juego y por fin a controlar ‘la historia facial’ (episódica y semántica) de un sistema, todo lo cual no es posible si se limita uno tan sólo a experimentar en el laboratorio con sujetos humanos.

En definitiva, las ciencias cognitivas implican el estudio teórico y experimental de los mecanismos de codificación, de almacenamiento, de manipulación y de transferencia de la información en sistemas naturales y artificiales que aseguran de manera integrada funciones de percepción, de razonamiento y de producción. Precisamente esta integración de funciones es lo que llamaré “conocimiento”. Pero mientras la psicología cognitiva se ocupa de las realizaciones de un sistema particular de conocimiento (el sistema nervioso central), las ciencias cognitivas tienen miras más comparativas y tratan de estudiar, más allá del carácter variable de las realizaciones sistemáticas, la “competencia” cognitiva misma.

\* Es decir, la generación estructural y funcional paso a paso de un proceso o de una conducta en un sistema físico de tratamiento de la información, lo cual es muy diferente de la simulación de un modelo por estimación del valor de los parámetros que entran en diversos juegos de ecuaciones.

### 3.2 Cognición y afecto

Por el momento las ciencias cognitivas tienen poco que decir sobre los afectos. Evidentemente no se trata de negar su existencia ni su importancia, aunque hay que reconocer que la descripción y explicación de los afectos plantean problemas extraordinariamente complejos. Por mi parte, no estoy lejos de considerar que el afecto constituye quizá una forma particular de “conocimiento”, sin duda inseparable de nuestro ambiente corporal, una experiencia que escapa por completo a los “puros” sistemas artificiales de tratamiento de la información. Una experiencia afectiva tiene indudablemente un contenido “representacional” de otras representaciones mentales. Para citar sólo un ejemplo, la sobreestimación teórica de la memoria semántica (conceptual) condujo a descuidar el papel de los afectos, que vuelven a adquirir toda su importancia cuando se estudian las propiedades de la memoria episódica (autobiográfica). Por lo demás, se sabe que las estructuras neurológicas que intervienen en el funcionamiento de la memoria (complejo hipocámpico, núcleos hipotalámicos y corteza frontal) suponen una elevada integración de sus componentes cognitivos y emocionales. En el nivel bioquímico, no existe una barrera insuperable entre el sistema nervioso de relación y el sistema nervioso vegetativo, de manera que puede predecirse el desarrollo próximo de una psiconeuroendocrinología que probablemente modifique nuestra concepción de las correlaciones entre estados fisiológicos emocionales, representaciones afectivas y representaciones cognitivas.

De todas maneras, cualesquiera que sean los problemas y debates epistemológicos engendrados por el concepto de cognición, me parece que los progresos decisivos no pueden realizarse en un solo terreno. Las ciencias cognitivas constituyen un “programa” de investigación, en el sentido que da al término Lakatos (1978), y su futuro dependerá de su capacidad heurística en el plano teórico y el plano experimental. Y todavía hace falta que se satisfaga un conjunto de condiciones institucionales mínimas. Desgraciadamente me temo que esto no ocurre todavía en nuestro país, pues nuestro atraso en este dominio es importante (tanto en el nivel de la formación como en el de la investigación). Es muy significativo el hecho de que

todavía estemos debatiendo sobre la naturaleza de lo cognitivo en nuestras sociedades científicas, siendo así que la Cognitive Science Society existe desde hace más de un decenio, que en Francia hace mucho tiempo que la ARC (Asociación para la investigación cognitiva) haya adoptado sus estatutos y que una Sociedad Europea de Psicología Cognitiva acaba de crearse muy recientemente (que yo sepa no existe todavía una sociedad europea de ciencia cognitiva). Cuando en el mundo existen numerosos centros de investigaciones de ciencias cognitivas, en el CNRS (Ciencias de la Vida, 1987), sólo he comprobado la existencia de cuatro laboratorios que en su título llevan el adjetivo “cognitivo” (sólo un laboratorio “se atreve a” llamarse de “Ciencias Cognitivas”). Y no hablemos de la formación en las ciencias cognitivas, que en Francia continúa estando todavía en estado embrionario cuando hay cursos completos (de seis años) o programas de doctorado establecidos hace muchos años en numerosos países (en los Estados Unidos, San Diego, Carnegie-Mellon, MIT, etc.; en Gran Bretaña, Edinburgo, Stirling, Brunel, etc.). En definitiva, el mejor medio de poder responder a la pregunta “¿Qué es lo cognitivo?” es sin duda procurar estudiar teórica y experimentalmente la manera en que percibimos nuestro ambiente, la manera en que lo comprendemos y la manera en que lo modificamos por nuestra acción. Esta empresa me parece hoy inseparable del programa científico de las Ciencias y Tecnologías del Conocimiento.

### Referencias bibliográficas

- Andler, D. (1987) (Guest Editor). Une nouvelle Science de l'Esprit: Intelligence Artificielle, sciences cognitives, nature du cerveau. *Le Débat*, 47.
- Arbib, M. A. (1987). Levels of modeling of mechanisms of visually guided behavior. *Behavioral and Brain Sciences*, 10, 407-465.
- Bonnet, C., Hoc, J.-M., y Tiberghien, G. (1986). *Psychologie, Intelligence Artificielle et Automatique*. Bruselas, Mardaga.
- Caverni, J.-P., Bastien, C., Mendelsohn y Tiberghien, G. (en prensa). *Psychologie cognitive: modèles et méthodes*. Grenoble, Presses Universitaires de Grenoble.
- De Mey, M. (1982). *The Cognitive Paradigm*. Dordrecht, Reidel.
- Fodor, J. A. (1983). *The Modularity of Mind: An Essay on Faculty Psychology*. Cambridge, MA, The MIT Press.
- Mandler, G. (1985). *Cognitive Psychology: An Essay in Cognitive Science*. Hillsdale, NJ, Lawrence Erlbaum.

- Squire, L. R. (1987). *Memory and Brain*. Nueva York, Oxford University Press.
- Tiberghien, G. (1985). Mais où sont les stimulus d'antan? *Psychologie Française*, 30, 177-184.
- Tiberghien, G. (1985). Fragmentes d'histoire de la psychologie. En J. Mathieu y R. Thomas, *Manuel de Psychologie*. París, Vigot.
- Tiberghien, G. (1986) (Guest Editor) Context and Cognition. *CPC: European Bulletin of Cognitive Psychology*, 6, 105-121.
- Tiberghien, G. (1986). Psychologie cognitive, sciences cognitives et technologies de la connaissance. En J.-L. Lemoigne (comp.). *Intelligence des mécanismes, Mécanismes de l'Intelligence*. París, Fayard.
- Tiberghien, G. (1986). Intelligence, mémoires et artifices. En C. Bonnet, J.-M. Hoc, y G. Tiberghien (comps.). *Psychologie, Intelligence Artificielle et Automatique*. Bruselas, Mardaga.
- Tiberghien, G. (en prensa). Face processing and face semantics. En A. W. Young y H. D. Ellis (comps.) *Handbook of Research on Face Processing*. Amsterdam, North Holland Publishing Company.
- Tiberghien, G. (en prensa). Symbolic and sub-symbolic models of human memory. En J. Delacour y J. Levy (comps.). *Systems with Learning and Memory Abilities*. Amsterdam, Elsevier & North Holland Publishing Company.
- Tiberghien, G. (en prensa). *Advances in Cognitives Science 2*. Chichester, Ellis Horwood.
- Winograd, T., Flores, F. (1986). *Understanding Computers and Cognition: A new foundation for design*. Norwood, NJ, Ablex.