



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO  
FACULTAD DE FILOSOFÍA Y LETRAS  
INSTITUTO DE INVESTIGACIONES FILOSÓFICAS



FACULTAD DE  
FILOSOFÍA Y LETRAS

LA DEFINICIÓN DE LOS ESTADOS MENTALES.  
UNA CRÍTICA A LAS TEORÍAS DE LA REDUCCIÓN  
EN LA FILOSOFÍA DE LA MENTE

TESIS  
QUE PARA OBTENER EL GRADO DE  
MAESTRO EN FILOSOFÍA DE LA CIENCIA  
PRESENTA

ADRIÁN ESPINOSA BARRIOS

DIRECTOR: DR. GUSTAVO ORTÍZ MILLÁN

SEPTIEMBRE DEL 2008





Universidad Nacional  
Autónoma de México



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

A Rosa María

A Norma y Julio, mis padres

A Alejandra y Alejandro

A mis amigos de generación

## AGRADECIMIENTOS

El presente trabajo es el producto final de dos años de labores en la Maestría en Filosofía de la Ciencia en el Instituto de Investigaciones Filosóficas de la UNAM. Agradezco al mencionado Instituto por la preparación durante este periodo. De igual manera, agradezco la ayuda de los profesores que directa o indirectamente intervinieron en la lectura, asesoría o corrección del trabajo. Muy especialmente quiero agradecer a los doctores Alfonso Arroyo, José Luis Díaz, Salma Saab, y Ángeles Eraña que participaron como sinodales en las revisiones, así como al doctor Gustavo Ortiz Millán por su tutoría como director del proyecto.

De igual manera, agradezco a la UNAM las facilidades otorgadas para la realización de mis estudios de posgrado y al CONACYT el apoyo económico recibido durante el periodo 2006-2008, sin los cuales no hubiera sido posible la presentación de este trabajo.

## ÍNDICE

<b>Introducción</b>	<b>4</b>
1. Estado físico y conducta, dos intentos de reducción	7
2. El funcionalismo computacional de Putnam	13
3. Dos versiones del funcionalismo	18
4. Los problemas de la definición funcionalista	21
a. Qualia: inversiones y ausencias, la primera objeción	23
b. Ceguera ante el significado, la segunda objeción	26
5. Un nuevo intento: la mente como adaptación	34
6. Los problemas de la mente como adaptación	40
7. ¿Cuál es el problema con las teorías de la identidad?	51
8. La reducción en la filosofía de la mente	55
<b>BIBLOGRAFÍA</b>	<b>55</b>

## INTRODUCCIÓN

Es tesis del presente trabajo que las reducciones ontológica y epistemológica en la filosofía de la mente, es decir, el intento por demostrar que los estados mentales (EM) *no son más que* entidades de un nivel teórico inferior o que tales estados pueden *ser explicados* por completo con base en conceptos de un nivel teórico inferior, no es la vía adecuada para conseguir una teoría completa de la mente. Esto se debe al compromiso de las teorías de la reducción con la suficiencia (ontológica o epistemológica) del nivel teórico al que presuntamente se reducen los EM. Tal compromiso plantea algunos de los problemas que se encuentran en el centro del debate en torno a la definición de los EM.

Las teorías de la reducción en la filosofía de la mente han sido de fuerte influencia durante el siglo XX; aquí se revisarán cuatro ejemplos: el fisicismo-tipo y el conductismo lógico por tratarse de dos casos de reducción ontológica y epistemológica respectivamente; el funcionalismo computacional en su primera versión de Hilary Putnam por tratarse del primer programa teórico de reducción epistemológica al rol funcional que resultó generador tanto de fuertes críticas como de nuevos intentos funcionalistas (se hace énfasis en la distinción entre dos versiones: los funcionalismos de primer y de segundo orden con el fin de ver los problemas que presenta cada uno frente al tema de la reducción); y la teoría "La-mente-como-adaptación" de Lawrence Shapiro por tratarse de un caso reciente de lo que es una

fuerte tendencia en la filosofía del mente: la naturalización de los EM<sup>1</sup> y porque el propio autor polemiza con el fisicismo-tipo y con el funcionalismo intentando corregir sus defectos. De tal modo, en los primeros 5 capítulos se exponen en líneas generales los postulados de cada una de las teorías revisadas así como sus principales críticas.

Creo, sin embargo, que el trabajo de las teorías de la reducción no es del todo desechable y que sus aportes serán necesarios en el estudio de los EM. Los aportes del estudio de la conducta, de las neurociencias, de los funcionalismos y de la biología evolucionista han de ser considerados en una eventual definición de los EM. No obstante, con el fin de ponernos en camino de una definición completa es necesario dejar de considerar estos aportes como necesarios y suficientes, lo que implica someter a revisión el programa reduccionista dentro de la filosofía de la mente. Creo que, de tal manera, es posible sentar las bases para iniciar el estudio de lo mental a partir de una perspectiva que no tenga como eje central el intento por reducir los EM a algún otro campo teórico previo, sino partiendo de la creación de un marco de referencia nuevo, con nuevos métodos que resulten acordes a las necesidades del objeto en cuestión. Los capítulos 6 y 7 están entonces dedicados a exponer algunas conclusiones sobre las teorías de la reducción revisadas previamente así como a hacer un análisis del tema de la reducción en la filosofía de la mente con miras a lograr eventualmente una ciencia completa de los EM.

---

<sup>1</sup> No se pretende mostrar con ello que todo intento naturalista de cualquier índole padezca de los mismos defectos que se critican en la teoría de Shapiro ni, mucho menos, que no sea posible trazar una teoría con base en la biología evolucionista que logre superar los problemas a los que se enfrentan las teorías aquí expuestas; simplemente se intenta mostrar que las teorías que pretenden explicar los EM con base en la suficiencia de algún elemento que explica por completo tales estados presentan problemas de difícil solución en virtud, precisamente, de su carácter reductor.

Las citas que se toman de la obra de Lawrence Shapiro son traducciones directas del inglés realizadas por mí. El resto de las citas son tomadas de versiones de los textos en español.



## 1. ESTADO FÍSICO Y CONDUCTA: DOS INTENTOS DE REDUCCIÓN

Uno de los problemas centrales de la filosofía de la mente contemporánea es el de la definición de los EM. A lo largo del siglo XX, diversas teorías han sido propuestas para resolver este problema<sup>2</sup>. Sin embargo, ninguna de estas opciones ha sido aceptada unívocamente y el debate, lejos de encontrarse superado, está hoy, a la vista de los avances en las llamadas ciencias cognitivas, más vivo que nunca.

Intentar definir los EM por medio de una definición de carácter enumerativo implica enunciar los elementos que les son constitutivos a todos esos estados. Cada una de las teorías que aborden el problema incluirá o excluirá elementos de tal definición según considere si estos resultan, o no, pertinentes para la identificación completa del término *estado mental*. Ahora bien, una teoría que se compromete con una definición de tal concepto se compromete también con una forma de clasificación de los tipos de estados mentales. De la definición del concepto *estado mental* como aquello que posee las características  $\{c_1 \dots c_n\}$  se puede desprender una clasificación de los distintos tipos de EM al indicar en qué forma han de variar  $c_1 \dots c_n$  para incluir a cada ejemplar en una de las categorías de la clasificación. Así, bajo la teoría  $T_1$ , los términos  $x_1, x_2, x_3$  que se usan para designar ciertos estados, pueden inscribirse dentro de la definición de EM propuesta por la teoría al cumplir con las condiciones generales  $EM = \{c_1 \dots c_n\}$ . Sin embargo,  $x_1, x_2, x_3$  pueden poseer a su vez diferencias

---

<sup>2</sup> Existe abundante bibliografía en donde se hacen revisiones de las distintas teorías propuestas para resolver el problema mente-cuerpo. *Vid.*, por ejemplo, J. Kim, *Philosophy of mind*; P. Churchland, *Matter and consciousness: A contemporary introduction to the philosophy of mind*; K. Campbell, *Cuerpo y mente*; J. Searle, *Mind: A brief introduction*; W. Lycan, "The Mind-Body Problem", *The blackwell guide to philosophy of mind*, Ch. 2, pp. 47-64.

específicas cuya identificación a partir de la definición general permite agruparlos en categorías más restringidas.

Hay un conjunto de teorías, que aquí llamaremos teorías de la reducción, que han tenido amplia influencia dentro de la filosofía de la mente del siglo XX. Siguiendo el modelo de la ciencia física, algunas teorías de la reducción intentan desvelar lo que en *realidad* se esconde detrás del concepto de EM.<sup>3</sup> Tal es el caso del fisicismo-tipo<sup>4</sup>. Una teoría que se inscriba dentro de los postulados del fisicismo-tipo daría una definición de los EM en los siguientes términos: EM = {Estimulación de una región cerebral C, de modo tal que al estado mental EM<sub>1</sub> le corresponde la estimulación de la región cerebral C<sub>1</sub> en todos los casos y sólo de ella}. De aquí se puede derivar una clasificación de los EM de acuerdo a la topografía cerebral: ante la estimulación de la región cerebral C<sub>1</sub> se presentan los estados mentales del *tipo* EM<sub>1</sub>, por lo tanto, dos estados EM<sub>α</sub> y EM<sub>β</sub> que se presenten al estimular la región cerebral C<sub>1</sub> caerán bajo la categoría EM<sub>1</sub>, es decir, serán del mismo *tipo*.<sup>5</sup>

---

<sup>3</sup> Son recurrentes las alusiones a los casos en que la física ha logrado hacer este desvelamiento con conceptos de especial utilidad: la ciencia actual nos ha demostrado que la temperatura, por ejemplo, *no es más que* la energía cinética media de las moléculas de un cuerpo o fluido, algo similar ocurre con casos como la propiedad del vidrio de ser frágil, o con la explicación de los rayos en una tormenta; *vid.* D. K. Lewis, "Psychophysical and theoretical identifications"; John Searle, *Mentes cerebros y ciencia*, Cap. 1; D.M. Armstrong, *The nature of mind*; J.J.C. Smart, "Sensations and brain processes". Más adelante regresaré al concepto de identidad, *vid infra* cap. 3 y cap. 7.

<sup>4</sup> Se utilizará el término "fiscicismo" para traducir el inglés "*physicalism*". Es traducción común en español el término "fisicalismo", sin embargo, puesto que no hay en nuestro idioma el adjetivo "físical", y sí por el contrario el adjetivo "físico", entonces se preferirá el término señalado (Observación del Dr. Gustavo Ortiz-Millán).

<sup>5</sup> Bajo el concepto de fisicismo se engloban aquellas teorías que afirman que la explicación de los términos mentales queda dada totalmente a partir de los principios físicos que explican el funcionamiento del sistema en que se realizan los EM (el cerebro). Esta descripción resulta simplificada para la gran variedad de fisicismos que hay. Sin embargo, la tomaremos aquí como ejemplo del intento de reducción de los EM a una instancia que puede ser estudiada por disciplinas como la física y las neurociencias. De los varios tipos de fisicismos, el más radical sería el presentado aquí como fisicismo-tipo (*type physicalism*) que postula que los EM no son otra cosa más que estados físicos del cerebro. Es este tipo de fisicismo el que sirve como punto de partida de las críticas formuladas por el

Otro intento de definición por reducción de los términos mentales fue el del conductismo lógico o filosófico. El conductismo lógico es, sin embargo, un intento diferente de reducción del fisicista; mientras el fisicismo afirma en general que los EM *no son más que* estados físicos, el conductismo lógico afirma que *lo que significan* los términos mentales que utilizamos *no es más que* conducta o disposiciones conductuales. En ese sentido, el conductismo lógico no se compromete con una tesis de carácter ontológico, sino que sólo hace una afirmación sobre la semántica de nuestro lenguaje mental. En última instancia, para el conductista lógico, el problema mente-cerebro es un pseudoproblema que se diluye una vez que reconocemos que no es a una instancia inmaterial (la mente) a lo que nos referimos cuando usamos nuestro vocabulario mental, sino sólo a conducta observable o posible (para el conductista lógico queda abierto el problema de decidir si tal instancia realmente existe). Entonces, para este tipo de conductismo un EM cualquiera se define, en forma completa y sin pérdida de significado, como: EM = {estímulos, disposiciones a actuar}. De modo que la taxonomía respectiva ordenará a los EM de acuerdo a los *tipos* de estímulos que reciba el sistema y al *tipo* de conducta que esté en disposición a manifestar.<sup>6</sup>

---

funcionalismo, es por ello que es el fisicismo que se toma en cuenta en este trabajo. Por otra parte, alguna variedad del funcionalismo puede ser considerado un tipo de fisicismo (*token physicalism*) *vid infra* cap. 3. Existen diversos tipos de compromisos fisicistas que se expresan en otras tantas teorías. La teoría de Smart ("Sensations and brain processes"), la superveniencia de lo mental a lo físico (J. Kim, *Mind and Supervenience*) y teorías como el monismo anómalo de Davidson (D. Davidson, "Eventos mentales", *Ensayos sobre acciones y sucesos*), pueden ser consideradas como ejemplos de teorías fisicistas. Para información sobre el fisicismo *vid.*: Daniel Stoljar, *Physicalism*.

<sup>6</sup> Al conductismo lógico también se le conoce como conductismo filosófico o analítico. Este tipo de conductismo, que es el que aquí nos interesa, tiene sus raíces, por un lado, en *The concept of mind* de G. Ryle y, por otro, en el programa verificacionista del positivismo lógico. No hay que confundir el conductismo lógico con el conductismo metodológico que intenta operacionalizar la conducta con fines de observación empírica. Para una revisión sobre el conductismo *vid.*, G. Graham, *Behaviorism*.

Tanto el fisicismo-tipo como el conductismo son teorías de la reducción, pues en ambos casos se intenta *identificar* o *entender* el concepto de EM por medio de otro conjunto de términos de los que puede dar cuenta la ciencia, ya sea la física o la ciencia conductual (psicología). Ambos postulados detrás de múltiples teorías fisicistas y conductistas se enfrentaron a fuertes objeciones que las desacreditaron como corpus teóricos válidos para dar una explicación completa de los EM. Sin embargo, hay que reconocer que ambas teorías sentaron las bases de la discusión ofreciendo la pauta para ulteriores propuestas.

El problema al que se enfrentan estos dos postulados teóricos es el mismo que se presentará a lo largo de todas las teorías de la reducción: para que el conductismo lógico y el fisicismo-tipo fueran ciertos, necesitarían demostrar que hay una relación de equivalencia entre el significado de los términos mentales y la conducta o entre EM y estados físicos, es decir, demostrar que la conducta, o los estados físicos, son condición necesaria y suficiente para la definición de los EM. En tal caso, a cada EM (el dolor, por ejemplo) le debería corresponder uno y sólo un estado conductual y viceversa, o que a cada estado mental le correspondiera uno y sólo un estado físico. Pero esto parece no ser tan fácil de demostrar. Así como hay una gran variedad de conductas reales y posibles que podrían ser consecuencia de un mismo estado mental, hay también varios estados mentales que podrían ser la causa de la misma conducta. Además, el conductismo presenta el problema de la presunta imposibilidad de eliminar por completo los términos mentales del vocabulario conductual. En efecto, la mente presenta una característica que se ha dado en llamar *holismo mental* que hace

que cada EM esté relacionado a su vez con otros tantos EM de modo que cada uno de ellos depende en forma directa o indirecta, del resto de estados del sistema. Esto provoca que el análisis en términos conductuales de un EM cualquiera (la esperanza de ganar la lotería, por ejemplo) implique el uso de otros términos mentales (la creencia de que hay una probabilidad, aunque pequeña, de ganar, el deseo de obtener el premio, etc.). Siempre parece surgir la necesidad de utilizar conceptos que van más allá de la pura conducta. Esto hace fracasar al programa que pensaba que para definir los términos mentales era *suficiente* el vocabulario conductual.

El problema de la falta de equivalencia se encuentra también presente al afirmar la identidad de los EM con los estados físicos. La principal crítica al fisicismo-tipo señala que bien podemos pensar en sistemas que tengan bases físicas distintas a la nuestra (otra organización neurológica) y que, sin embargo, posean EM. Dentro de esta objeción se inscriben los intentos de adscribir mentalidad tanto a otras especies distintas de la humana, como a eventuales seres artificiales o extraterrestres. A este tipo de argumento contra la reducción del fisicismo-tipo se le conoce como argumento de la *múltiple realizabilidad* de lo mental. Por otro lado, la equivalencia también se ve puesta en duda cuando los estudios neurofisiológicos demuestran que las mismas regiones cerebrales pueden instanciar múltiples funciones mentales.<sup>7</sup> Si ambas

---

<sup>7</sup> Los estudios neurológicos actuales sobre la plasticidad cerebral podrían servir de constatación empírica del argumento de la múltiple realizabilidad al demostrar que ciertas estructuras cerebrales no siempre son necesarias ni suficientes para la realización de cierta función mental. No habría, por lo tanto, identidad entre EM y estimulación de una región cerebral en particular. Existe múltiple bibliografía al respecto, *Vid* F. Boller, "Rational basis of rehabilitation following cerebral lesions: A review of the concept of cerebral plasticity"; E. Bifill y E. Carbonell, "Conducta simbólica y neuroplasticidad: ¿Un ejemplo de coevolución gen-cultura?", J. Nithianantharajha y A. J. Hannan, "Enriched environments, experience-dependent plasticity and disorders of the nervous system"; R. Lewin, "Is your brain really necessary?".

objecciones resultan viables, entonces los campos teóricos “mental” y “físico” no pueden corresponderse unívocamente.

Estos argumentos que tienden a demostrar la insuficiencia del postulado conductista y del fisicista para la definición de los EM fueron el inicio de la construcción de una nueva teoría que tendría como objetivo salvar estos obstáculos: el funcionalismo. Sin embargo, el intento resultó fallido, pues el propio funcionalismo, a la postre, evidenciaría los mismos errores que sus predecesoras. En el siguiente apartado expondremos la teoría del funcionalismo computacional como fue originalmente postulada por Putnam para, posteriormente, hacer algunas distinciones dentro de la propia teoría funcionalista.<sup>8</sup>

---

<sup>8</sup> El propio Putnam abandonaría con el tiempo su propia teoría. *Vid.* H. Putnam, *Realidad y representación*, sin embargo, sigue siendo pertinente el repaso de las críticas al funcionalismo por dos razones: 1) el funcionalismo siguió siendo un programa teórico desarrollado más allá de sus precursores (J. Kim, D. K. Lewis, los Churchland, S. Shoemaker, etc.) y 2) Nuevas teorías, como la de Shapiro, pretenden rescatar lo mejor del funcionalismo evitando sus problemas.

## 2. EL FUNCIONALISMO COMPUTACIONAL DE PUTNAM

Frente a los reduccionismos conductistas y físicos, Hilary Putnam presentó una tercera teoría de los EM<sup>9</sup>: el funcionalismo computacional<sup>10</sup>. La afirmación es que los EM *no son otra cosa más* que estados funcionales. ¿Qué son los estados funcionales? Para responder a esta pregunta es necesario introducir el concepto de máquina de Turing en el cual se basa Putnam para proponer el concepto de autómata probabilístico. Una máquina de Turing es un dispositivo idealizado creado por Alan Turing en la cuarta década del siglo XX.<sup>11</sup> En el dispositivo de Turing se pueden introducir datos (*input*) para obtener ciertos resultados (*output*) de acuerdo a las instrucciones de la máquina y a la situación interna (estados internos) en que se encuentra. Este dispositivo es totalmente determinista en el sentido de que dado cierto *input* introducido al sistema y el estado interno en que éste se encuentre, sabremos siempre con absoluta certeza cuál será el *output* obtenido y el estado interno subsiguiente. Además, se trata de un dispositivo que afirma relaciones causales: los *inputs* causan que el dispositivo se encuentre en determinado estado, el cual, a su vez, es causa de los *outputs* que el dispositivo ha de emitir dadas las relaciones en que dicho estado se encuentre respecto de otros estados del sistema. La transición entre

---

<sup>9</sup> H. Putnam, *Cerebro y conducta*; *id.* *La vida mental de algunas máquinas*; *id.* *La naturaleza de los estados mentales*; *id.* *Otras mentes*.

<sup>10</sup> El funcionalismo ha sido un programa teórico muy fecundo, de modo que cuenta también con diversas variantes. Los funcionalismos que mencionaré en este trabajo serán el computacional y el analítico (de segundo y primer orden, respectivamente). Será necesario mantener presente esta aclaración toda vez que, por un lado, no todos los tipos de funcionalismo son susceptibles de las mismas críticas y, por otro, es posible rescatar algunas de las ideas funcionalistas sin que por ello se pueda inscribir tal intento dentro del funcionalismo computacional o del analítico. Para una revisión del funcionalismo *vid.* N. Block, "Troubles with functionalism", *Readings in the Philosophy of Psychology*, vol. 1.; J. Levin, *Functionalism*.

<sup>11</sup> *Vid.* A. Turing, "On Computable Numbers, With an Application to the Entscheidungsproblem"; R. Penrose, *La mente nueva del emperador*, pp. 53 y ss.

estados internos viene dada por un algoritmo que técnicamente se llama *tabla de máquina*. Así, una máquina de Turing puede tener una tabla de máquina que le indique: si recibes el input  $I_1$  y te encuentras en el estado interno  $E_1$ , entonces emites el resultado  $O_1$  y pasas al estado interno  $E_2$ .

Putnam toma la idea de una máquina de Turing<sup>12</sup> y la modifica para proponer un modelo bajo el cual cabría pensar cómo funciona la mente humana. A este modelo le llama un autómata probabilístico. Cabe suponer que la mente humana funciona bajo principios probabilísticos. Putnam parece entender por esto dos cosas distintas: 1) la mente puede evaluar probabilidades de estados de cosas y actuar maximizando la utilidad de acuerdo a la teoría racional de la decisión y 2) las respuestas conductuales y los estados mentales de una persona, aunque los podamos inferir o predecir con cierta certeza, se encuentran indeterminados. Así pues, un modelo computacional que se ajuste a la mente humana será una máquina de Turing modificada: un autómata probabilístico. Este autómata estará provisto de órganos sensoriales para el ingreso de información (input), órganos motrices (output) y estados internos cuya transición esté especificada por un algoritmo probabilístico del tipo: si recibes el input  $I_1$  y te encuentras en el estado interno  $E_1$ , entonces *prefiere* (con un determinado grado de probabilidad) la respuesta  $O_1$  y *prefiere* (con un determinado grado de probabilidad) pasar al estado  $E_2$ . Los estados internos  $E$  son entonces estados que se definen de acuerdo a su relación probabilística causal con los datos de entrada, los datos de salida y de otros estados. Es decir, para cada estado  $E_n$  hay un conjunto, y sólo uno,  $\{I_n, O_n, R_n\}$  que lo define, en donde  $I$  son los datos de entrada,  $O$  los datos de salida y  $R$  las

---

<sup>12</sup> H. Putnam, *La vida mental de algunas máquinas*.



relaciones causales que mantiene con otros estados.  $E_n$  queda entonces definido como un estado funcional del autómata probabilístico.

A partir de esta caracterización del funcionalismo es fácil ver que un autómata como el descrito puede ser instanciado en un organismo biológico (como el ser humano), en algún otro sistema físico (una computadora) o, incluso, un sistema no físico. En caso de que la analogía con la mente humana sea correcta, esto equivale a decir que los estados mentales *son* estados funcionales. Un dolor, como ejemplo de EM, queda definido funcionalmente como un estado del sistema *normalmente* ocasionado por un daño ocurrido en su estructura física (input), que da origen a ciertos otros estados internos (creencia de que se está sufriendo un daño, deseo de retirar la parte afectada, etc.) y a la *inclinación* espontánea a retirar la parte afectada (output) de la fuente que causa el dolor.

Es necesario aquí aclarar algunos puntos importantes sobre este funcionalismo expuesto que serán relevantes para la discusión posterior.

1. El funcionalismo computacional, a diferencia del conductismo lógico, admite la existencia de estados internos del sistema y de sus poderes causales. El hecho de que el sistema exhiba determinada conducta es consecuencia de la existencia de cierto estado interno que lo produce; de tal modo es explicable que existan casos en que el sistema no exhiba una conducta normalmente asociada con determinado estado y, no obstante, estemos dispuestos a adjudicar la posesión de tal estado por

parte del sistema.<sup>13</sup> En esto consiste parte de la crítica del funcionalismo computacional contra el conductismo lógico: los EM no son disposiciones conductuales, en realidad, son causa de tales disposiciones.

2. Aunque el funcionalismo es compatible con el fisicismo, en el sentido de que se da por sentado que los dispositivos físicos que instancian autómatas probabilísticos están sometidos a las leyes físicas, sin embargo el funcionalismo computacional de Putnam *no es* un fisicismo en el mismo sentido del fisicismo-tipo. Y no lo es debido a la tesis de la múltiple realizabilidad de los estados funcionales. Lo que importa para caracterizar los EM son las relaciones que éstos guardan con otros EM así como con los datos de entrada y de salida; es decir, lo que importa es su rol funcional y no la base material en que se realizan. De hecho, el funcionalismo se presenta como una refutación al fisicismo-tipo en dos sentidos: a) dar una descripción completa de un EM, afirma el funcionalismo, es dar su estado funcional y esto no dice nada sobre su estado físico y b) como vimos, para que el fisicismo-tipo sea cierto, es necesario que se dé la equivalencia: "Hay un estado mental  $x$  si y sólo si hay estimulación de las fibras  $C$ "; sin embargo, el funcionalismo afirma que un EM (el dolor, por ejemplo) considerado como un estado funcional puede ser realizado en múltiples dispositivos físicos diferentes que no necesariamente han de corresponderse con las fibras  $C$ . Se puede, entonces, inferir que dar una definición funcional de un EM implica especificar *todas* sus relaciones funcionalmente pertinentes (con otros EM, con datos de entrada y de salida), pero entre estas relaciones causales no figurarán aquellas que el EM mantenga con el dispositivo físico en que se realiza. De esta manera, el

---

<sup>13</sup> Vid, el ejemplo de los superespartanos elaborado por Putnam, *Cerebro y conducta*.

funcionalismo supone que la instancia física no es un elemento necesario a incluir dentro de la definición de los EM.

3. Putnam señala explícitamente<sup>14</sup> que la afirmación de la analogía (o eventual identidad) existente entre la mente humana y un autómata probabilístico no pasa de ser una hipótesis verosímil (mucho más, quizá, que las hipótesis materialista y conductista) que ha de ser comprobada o refutada empíricamente, pero que no se trata en absoluto de un argumento deductivo contundente. Y esto es así por dos razones: a) aunque parezca bastante verosímil que la mente humana funciona bajo leyes probabilísticas, también parece cierto que en las situaciones cotidianas el ser humano no actúa siempre de acuerdo a la teoría de la acción racional, además, b) el modelo esbozado de un autómata probabilístico parece ser en extremo simplificado y racional como para dar cuenta de la enorme complejidad de la mente humana. Sin embargo, dice Putnam, esto último no es por sí mismo un defecto, pues el modelo podría ser complejizado lo suficiente como para semejar lo más posible a la mente humana, pero, por lo pronto, basta con que dicha mente se comporte en términos generales como lo indica el modelo para que las conclusiones extraídas sean verosímiles. Será necesario esperar a la elaboración de una teoría psicológica desarrollada que sea capaz de mostrar que los EM se pueden definir por completo por medio de su rol funcional.

---

<sup>14</sup> H. Putnam, *La vida mental de algunas máquinas; id., La naturaleza de los estados mentales.*

### 3. DOS VERSIONES DEL FUNCIONALISMO

Actualmente, existen múltiples variedades de funcionalismo así como clasificaciones que intentan agruparlas según rasgos en común.<sup>15</sup> Aquí haré mención de la clasificación de García Carpintero que distingue entre funcionalismos de primer y de segundo orden por considerar que tal distinción deja en evidencia la forma en que enfrentan los problemas de la reducción las dos variedades mencionadas.<sup>16</sup>

Como se mencionó arriba<sup>17</sup>, una imagen recurrentemente usada en las teorías de la reducción son las identificaciones llevadas a cabo en la ciencia. La forma general de tales identificaciones es la siguiente: un fenómeno macroscópico que era anteriormente explicado en términos fenoménicos (la temperatura, el color, etc.) se explica ahora recurriendo a su estructura microfísica (energía cinética media, amplitud de onda, etc.). De tal forma operan las identificaciones teóricas en la ciencia. El funcionalismo de primer orden, siguiendo esta estructura, afirma algo similar respecto de los términos mentales. Para esta versión funcionalista hay una identificación teórica entre la definición funcional de un EM y la instancia física que lo realiza. Tal identificación es posible gracias a que los EM se encuentran definidos, según el funcionalismo de primer orden, de la misma manera que son definidos los términos científicos como “masa”, “peso”, “gen”, es decir, por medio del conjunto de relaciones que mantienen con el resto de los términos dentro de la teoría. Es gracias a esta identificación que los EM poseen toda su eficacia causal; a final de cuentas no son

---

<sup>15</sup> Vid por ejemplo, S. Shoemaker, “Some varieties of functionalism”.

<sup>16</sup> M. García Carpintero, “El funcionalismo”.

<sup>17</sup> Vid *supra* n. 3.

otra cosa más que estados físicos. En este sentido, el funcionalismo de primer orden puede ser entendido como un caso de fisicismo-tipo pues lo que afirma es que, una vez que logramos apreciar las relaciones funcionales en que se encuentra un EM, nos damos cuenta de que tales relaciones no son sino estados físicos. Las propiedades funcionales sólo están al nivel de las relaciones interteóricas de los conceptos, pero por sí mismas, no forman un campo irreductible a la base física.<sup>18</sup>

Por otra parte, el funcionalismo de segundo orden tiene como centro de su argumentación la tesis de la múltiple realizabilidad. Para este tipo de funcionalismo, los EM no son más que roles funcionales, pero estos, a diferencia de lo que afirma el funcionalismo de primer grado, no son roles identificables con una instancia física determinada. De hecho, es perfectamente posible que se presente el mismo rol funcional en distintas realizaciones físicas. El sustrato físico que subyace al rol funcional se encuentra en un nivel de primer orden; el rol funcional se presenta sobre el primer nivel pero, por sí mismo está en un nivel que no puede ser identificado con ninguna de las instancias físicas en las que es posible su realización. En otras palabras, entre dos sistemas que presenten el mismo rol funcional la estructura microfísica puede ser distinta en un amplio margen. Como se puede apreciar, el funcionalismo computacional de Putnam es el principal representante de esta versión.<sup>19</sup>

---

<sup>18</sup> García Carpintero señala al funcionalismo analítico de D. Lewis como el mayor representante de esta variedad funcionalista; *vid* D. Lewis, "How to define theoretical terms"; *id.* "Mad pain and martial pain".

<sup>19</sup> *Vid* H. Putnam, *Cerebro y conducta*; *id.* *La vida mental de algunas máquinas*; *id.* *La naturaleza de los estados mentales*; *id.* *Otras mentes*. Según García Carpintero, esta es la versión del funcionalismo más popular; la mayoría de los funcionalistas acepta la tesis de la múltiple realizabilidad desde su presentación por Putnam. Es precisamente esta relación entre los niveles de primer y de segundo orden la que está en el centro de la tesis funcionalista de la superveniencia de Kim. *Vid* J. Kim, *Mind and supervenience*.

Estas dos versiones del funcionalismo enfrentarán el tema de la reducción de formas distintas. Para el funcionalismo de primer orden es claro que una vez que se definen los términos mentales funcionalmente es posible llevar a cabo la reducción al nivel físico. En cambio, aceptar la tesis de la múltiple realizabilidad de lo mental implica aceptar que la instancia física no es pertinente para definir el rol funcional. Rol funcional e instancia física no son lo mismo, por lo tanto no hay reducción de lo mental a lo físico. De este modo, podemos ver que hay cuando menos dos posturas funcionalistas en torno a la reducción: un funcionalismo reductor de lo mental a lo físico y otro que afirma que tal reducción no es posible.

Ambos tipos de funcionalismo enfrentan diferentes problemas. La mayoría de las críticas intentan hacer inválida la equivalencia entre EM y rol funcional o entre rol funcional e instancia física. Las críticas a las que está expuesto el funcionalismo de primer orden son similares a las que enfrentan los fisicismos-tipo o exceden el tema de la reducción.<sup>20</sup> Es por ello que a continuación nos enfocaremos en las críticas que van en contra de la afirmación de los funcionalismos de segundo orden de que los EM son roles funcionales.

---

<sup>20</sup> Un ejemplo de crítica al funcionalismo de primer orden que no tiene relación directa con el tema de la reducción es el del problema de considerar a los EM y sus relaciones como una teoría *popular* de la mente.

#### 4. LOS PROBLEMAS DE LA DEFINICIÓN FUNCIONALISTA

La propuesta del funcionalismo computacional de los EM sonaba prometedora al intentar avanzar entre los dos riesgos que presentaban las teorías fisicista y conductista: el Escila de la estrechez fisicista y el Caribdis de la indeterminación conductista. El antídoto ante ambos riesgos era la aportación propiamente funcionalista: las relaciones causales de los estados internos del sistema, es decir, el rol funcional. Para poder solucionar los problemas a los que se enfrentaban el conductismo lógico y el fisicismo-tipo es necesario demostrar que los EM son roles funcionales. Es esta la afirmación que no comparten los críticos del funcionalismo. En este capítulo revisaremos dos argumentos que se han dado contra el funcionalismo computacional para demostrar que éste no logra establecer la equivalencia entre rol funcional y EM. Lo que dicen los críticos, en términos generales, es que quizá el rol funcional es múltiplemente realizable, pero que los EM no son roles funcionales.

Intuitivamente, los EM se nos presentan con dos características que parecen inherentes a ellos: a) los EM se “sienten” de una forma determinada para quien los experimenta y b) los EM se “refieren” a algo fuera de ellos mismos. Un dolor en el pie, la percepción de rojo o un estado de ira se “sienten” de forma distinta y mi conocimiento de que el sol es una estrella, mi creencia de que el libro está sobre la mesa o mi deseo de que mi boleto de la lotería salga premiado, se “refieren” a distintas cosas. A ese componente subjetivo y fenomenológico que señala la forma en que se sienten los EM se le conoce dentro de la filosofía de la mente como *qualia*. Si intuitivamente pensamos que la forma en que es sentido un dolor es parte inherente

del dolor en sí mismo, entonces esperaríamos que una teoría aceptable de los EM logre capturar los qualia dentro de la definición.<sup>21</sup> Por otro lado, al tema de la referencia de los EM se le conoce, desde Brentano, como el problema de la *intencionalidad*<sup>22</sup>.

Es un debate abierto si todos los EM presentan ambas características o más bien resulta ser cada una de ellas la característica definitoria de un subgrupo dentro de los EM: estados cualitativos, en un caso; estados intencionales, en el otro. De cualquier modo, para que los temas de los qualia y de la intencionalidad representen problemas a la definición del funcionalismo de segundo orden, son necesarias dos cosas: 1) la afirmación de que se trata en ambos casos de características inherentes a los EM y, por lo tanto, que son elementos necesarios a integrar en la definición y 2) la demostración de que es posible tener estados funcionalmente equivalentes a EM que no presenten qualia y/o intencionalidad, con lo cual se demostraría que la pretendida identidad entre EM y rol funcional no es válida. A continuación se exponen dos objeciones que tienen la finalidad de demostrar la incompletud de la definición del funcionalismo computacional.

---

<sup>21</sup> El tema de los qualia ocupa un amplio espacio en la discusión de la filosofía de la mente, además del texto de Fodor y Block del que se hablará adelante, hay muchas otras referencias clásicas; *vid* T. Nagel, "¿Cómo es ser un murciélago?", F. Jackson, "Lo que María no sabía", D. Dennett, "Quining Qualia", etc.

<sup>22</sup> *Vid.*, F. Brentano, *Psicología desde un punto de vista empírico*.



### a. Qualia: inversiones y ausencias, la primera objeción

Es posible que cuando hacemos un análisis en primera persona de nuestros propios EM podamos distinguir en ellos los elementos causales de entrada, las conductas que son provocadas por la existencia de tal estado y las relaciones que éste mantiene con otros EM, sin embargo, cada uno se puede percatar que en tal análisis existe un elemento de, cuando menos, algunos EM que no corresponde a ninguno de los elementos de los que nos habla la definición funcionalista: los qualia. Lo problemático del caso es que tal característica de ciertos EM parece no ser sólo accesoria, sino esencial. En efecto, la forma en que siento un dolor o la forma en que percibo un color parece ser vital para esa sensación o para esa percepción, sin ella, parecería no haber dolor ni experiencia de color. El problema radica en que, a diferencia de ciertos EM que parecen poder ser totalmente definidos por sus relaciones causales funcionalmente pertinentes (a veces llamados estados proposicionales), en los estados en los que sin duda se presenta el rasgo cualitativo hay una característica que parece esencial y que escapa de la definición funcional.<sup>23</sup>

Quienes pusieron el acento en este problema que presentaba el funcionalismo computacional fueron Ned Block y Jerry Fodor en *Lo que no son los estados*

---

<sup>23</sup> Al espacio que se abre entre la experiencia fenoménica y la integración de dicha experiencia en una teoría científica de la mente se le conoce como el problema de la *brecha explicativa*. Más allá del funcionalismo, hay cuando menos un intento, el de Dennett, por crear un método que logre cerrar la brecha haciendo accesible la experiencia subjetiva para terceras personas por medio de los reportes verbales; *vid infra* p. 49.

psicológicos<sup>24</sup>. Ahí, se presentan una serie de dificultades a las que se enfrenta un funcionalismo como el de Putnam al que ellos definen como una “Teoría de la Identidad de Estados Funcionales” (TIEF) y del cual pretenden demostrar que presenta una incorrecta definición de los EM al no reconocer como rasgo esencial de éstos su carácter cualitativo. De las seis objeciones presentadas en el artículo, cinco de ellas se dirigen contra la identificación de los EM con estados funcionales. Sin embargo hay una de ellas, la de los qualia ausentes y el espectro invertido (que los autores tratan como una sola objeción), que de ser cierta atacaría al funcionalismo en general al afirmar que no sólo no es posible la identificación deseada, sino que la identidad o diferencia entre dos EM no puede ser establecida *en absoluto* “en términos funcionalistas y que no pueden corresponder a ninguna propiedad definible en términos de autómatas abstractos”<sup>25</sup>. Esto es tanto como afirmar que la definición derivada de las tesis funcionalistas no sólo es liberal, restrictiva o incompleta (argumentos con los que se criticaba a los fisicismos y a los conductismos), sino completamente errónea. He aquí las dos objeciones relacionadas con los qualia.

**Hipótesis del Espectro Invertido (HEI):** La crítica de Block y Fodor es que la TIEF deja fuera de consideración un elemento indispensable de, cuando menos, algunos EM: los qualia. Esto se demuestra de la siguiente manera: es posible pensar en dos casos en los que cambie el estado cualitativo permaneciendo, no obstante, idéntico el estado funcional de ambos. Dos EM (en la misma persona o en personas distintas) pueden ser idénticos funcionalmente a pesar de que sus respectivos qualia difieran

---

<sup>24</sup> N. Block y J. Fodor, “Lo que no son los estados psicológicos”, *La naturaleza de la experiencia. Vol. I Sensaciones*, pp. 155-181. Para una exposición más clara y completa de las objeciones al funcionalismo véase N. Block, “Troubles with Functionalism”, *Readings in the Philosophy of Psychology*, pp. 268-305

<sup>25</sup> N. Block y J. Fodor, *op. cit.* P. 171 (énfasis mío).

ligera o radicalmente. Intuitivamente estaríamos tentados a decir que dos EM de dolor que “se sientan” diferente o dos percepciones de color que “se aparecen como” diferentes a dos sujetos (o al mismo sujeto en diferentes momentos) son distintos EM a pesar de estar causados por los mismos datos de entrada, causar las mismas acciones de salida y encontrarse en la misma relación con otros EM. Tales ejemplos de inversiones son lógicamente posibles (no causa contradicción pensar en su existencia), pero, además, parece que los distintos casos de ceguera ante el color dan evidencia empírica al respecto.<sup>26</sup>

**Hipótesis de los Qualia Ausentes (HQA):** Pero si efectivamente el carácter cualitativo carece de relevancia para la definición funcional, entonces se abre la puerta para una objeción más radical: dos EM podrían ser idénticos aun cuando uno de ellos posea carácter cualitativo y el otro no. Esta es la objeción que, según Block y Fodor, demostraría que “los estados psicológicos no pueden definirse en absoluto en términos funcionalistas”. De ser cierto el funcionalismo, dice Block en “Troubles with functionalism”, sería posible pensar en casos que parecen aberrantes a nuestras más claras intuiciones. Podríamos pensar en máquinas expendedoras con estados funcionalmente idénticos a nuestros EM, en autómatas dirigidos por homúnculos que hacen las veces de neuronas o en una población suficientemente grande (la de China, por ejemplo) en donde cada uno de sus miembros responda ante un dato de entrada produciendo un movimiento en un cuerpo artificial por medio de emisiones de radio. Nadie parecería dispuesto a pensar que la máquina expendedora tiene EM, que la población de China cumple las veces de un cerebro y que realiza EM o que un

---

<sup>26</sup> Vid A. Byrne y D. R. Hilbert, “How do things look to the color blind?”

autómata habitado por homúnculos tendría vida mental en el mismo sentido en que lo decimos de cualquier persona.

Block y Fodor no desarrollan más la objeción ni las posibles respuestas a la misma.<sup>27</sup> Desde el punto de vista de los autores, las dos hipótesis expuestas desacreditan por completo al funcionalismo como posible teoría de la mente, sin embargo, hacia el final de este capítulo aclararemos un par de puntos con la finalidad de comprender la crítica relacionada con la existencia de los qualia en su correcta dimensión.

#### **b. Ceguera ante el significado, la segunda objeción**

El segundo problema que ofrece la definición funcionalista de los EM es el de la *intencionalidad*. En 1874, Franz Brentano rescató este antiguo término escolástico para dar cuenta de una característica que él consideró propia de lo mental. Con este término, Brentano intentaba señalar a la propiedad de los fenómenos mentales de “dirigirse”, “apuntar” o “referirse” a objetos o estados de cosas.

¿Qué carácter positivo podremos, pues, indicar? ¿O acaso no hay ninguna definición positiva, que valga conjuntamente para todos los fenómenos psíquicos? [...] ya los antiguos psicólogos han llamado la atención sobre una especial afinidad y analogía que existe entre todos los fenómenos psíquicos, y en la que los fenómenos físicos no tienen parte. Todo fenómeno psíquico está caracterizado por lo que los escolásticos de la Edad Media han llamado la *inexistencia intencional* (o mental) de un objeto, y que nosotros llamaríamos, si bien con expresiones no enteramente inequívocas, *la referencia a un contenido, la dirección hacia un objeto (por el cual no hay que entender aquí una realidad), o la objetividad inmanente*. Todo fenómeno psíquico contiene en sí algo como su objeto, si bien

---

<sup>27</sup> La discusión, sin embargo, la continúan Block y Shoemaker en otra serie de famosos textos; *vid* N. Block, “Are absent qualia impossible?”; S. Shoemaker, “Funcionalismo y qualia”; *id.* “Absent qualia are impossible-A reply to Block”.

no todos del mismo modo. En la representación hay algo representado; en el juicio hay algo admitido o rechazado; en el amor, amado; en el odio, odiado; en el apetito, apetecido, etc. Esta inexistencia intencional es exclusivamente propia de los fenómenos psíquicos. Ningún fenómeno físico ofrece nada semejante.<sup>28</sup>

De tal forma, EM como las creencias, los deseos o las esperanzas son intencionales al poderse distinguir claramente los actos de creer, esperar o desear de aquello que en cada caso es creído, deseado o esperado. A este tipo de EM se les acostumbra llamar proposicionales por encontrarse dirigidos hacia una proposición que da cuenta del estado de cosas creído, deseado o esperado, y, en español, se encuentran normalmente expresados en oraciones con la partícula relativa “que”. Para Brentano era claro que la intencionalidad era la marca de lo mental, pues para él *todos* los fenómenos mentales, y *sólo* ellos, tenían la propiedad de referirse a objetos.<sup>29</sup>

La tesis de Brentano, como se le conoce, no está exenta de debate. Gran parte de la filosofía de la mente del s. XX ha puesto en duda que, en efecto, la intencionalidad sea la propiedad intrínseca de los EM. He aquí tres de los puntos conflictivos:

1. No resulta obvio que la intencionalidad sea una propiedad *exclusiva* de los fenómenos mentales; pareciera que algunos procesos físicos o biológicos (como los movimientos de algunas plantas en su búsqueda por la luz solar) pudieran ser

---

<sup>28</sup> F. Brentano, *Op. Cit.*, p. 81-82 (el énfasis es mío). Véase también: D. Jaquette, “Brentano’s concept of intentionality”, *The Cambridge Companion To Brentano*, p. 98.

<sup>29</sup> Brentano usa la expresión “inexistencia intencional” para los objetos a los que se dirigen los fenómenos mentales. Según Tim Crane, con esta expresión se quería dar a entender no sólo que los objetos o estados de cosas a los que se dirige el fenómeno mental podrían no existir –como en el caso de las creencias falsas, de las esperanzas o de los deseos– sino en el sentido de que el objeto al cual se dirigen podría encontrarse dentro de la mente misma. *Vid* T. Crane, “Intentionality as the mark of the mental”, *Contemporary Issues in the philosophy of mind*.

redescritos en términos intencionales sin la participación de alguna forma de fenómeno mental.

2. Más aún, parece que la intencionalidad no es ni siquiera característica propia de *todos* los fenómenos mentales. Aunque hay estados claramente intencionales (como los proposicionales), hay otros en los que parece absurdo buscar un objeto al cual se dirigen; tal es el caso de las sensaciones (dolor, comezón, ardor, placer) y de algunas emociones (algunas formas de ansiedad o de felicidad).<sup>30</sup>
3. ¿Qué quiere decir exactamente que un fenómeno mental esté *dirigido* a un objeto o estado de cosas?, ¿Qué pasa en casos, como en algunos deseos, en que el pretendido objeto no existe o no puede existir?

Los puntos 1 y 2 señalan respectivamente hacia la no suficiencia y la no necesidad de la intencionalidad para definir los EM. Sin embargo, aún es posible que aceptemos 1 y 2 y, al mismo tiempo, afirmemos que, si bien no en todos los EM, sí es cierto que en algunos de ellos la intencionalidad juega un papel primordial. Tal es la postura de John Searle para quien el significado de las palabras es un tipo de intencionalidad. En este caso, tenemos que los EM que intervienen en la comprensión y uso de un lenguaje presentan la relación intencional de la significación (afirmación que podría ser ofrecida como respuesta al problema 3)<sup>31</sup>.

---

<sup>30</sup> Sin embargo, incluso esta afirmación sobre la existencia de EM sin intencionalidad es objetable. Para una postura a favor de la existencia de EM en los que no hay objeto intencional véase J. Searle, *Intentionality*. Para la defensa de la postura de Brentano de la intencionalidad como la marca de lo mental, véase T. Crane, *op. cit.*

<sup>31</sup> Uno puede pensar que cada vez que un hablante hace uso de un lenguaje, i. e., cada vez que entiende y emite expresiones significativas, presenta actitudes proposicionales, pues ha de poseer ciertas creencias (sobre aquello a lo que las palabras *refieren*), deseos (de ser entendido por los interlocutores), esperanzas (de causar ciertas reacciones con sus emisiones), etc. Entonces, que un hablante entienda un lenguaje querría decir que establece una relación direccional entre las palabras u oraciones y los objetos

**El argumento del cuarto chino:** Con base en la afirmación de que en el uso y comprensión de un lenguaje intervienen EM intencionales, Searle elabora un *Gedankenexperiment*, el del cuarto chino, que pretendidamente ha de servir como contra-argumento de la Inteligencia Artificial fuerte, pero también del funcionalismo computacional. La idea es que, si se logra mostrar la posibilidad de crear un sistema que reproduzca las relaciones funcionales en que se encuentra una persona mientras usa y comprende un lenguaje y, sin embargo, no se puede decir de tal sistema que *comprenda* los signos que usa (toda vez que “comprender” sería conocer la relación intencional entre el signo usado y el objeto representado), entonces se habrá demostrado que el funcionalismo computacional no es una teoría válida para la explicación de los EM.<sup>32</sup>

El experimento mental es el siguiente: Una persona no-hablante del chino se instala en un cuarto con dos puertas, una de ellas sirve como *entrada* de signos chinos que le son enviados a la persona desde el exterior en tanto que la otra puerta ha de servir como *salida* de otros tantos signos chinos que el habitante del cuarto ordena y despacha con base en una tabla de instrucciones (escrita en su idioma) que correlaciona lo que para él son sólo figuras ininteligibles. La tabla de instrucciones está especialmente diseñada para que el habitante del cuarto pueda formar cadenas de

---

o estados de cosas que éstas representan, es decir que el hablante tiene la capacidad de entender a las palabras como *signos*.

<sup>32</sup> Vid J. Searle, “Minds, Brains and programs”, *The behavioral and brain sciences*, pp. 417-457. Aunque el artículo de Searle está dirigido principalmente contra la pretensión de la IA fuerte (como él la llama) de poder crear un sistema computacional que no sólo imite la comprensión lingüística, sino que efectivamente posea tal comprensión, sin embargo, sus conclusiones se pueden extender al funcionalismo computacional toda vez que él mismo afirma que tal teoría se encuentra a la base del proyecto de la IA fuerte.

figuras que una vez despachadas resultan significativas para la persona que las lea (y que *comprenda* el chino). Se puede observar que un sistema tal reproduce el estado funcional de un hablante de chino y, no obstante, afirma Searle, ni el sistema en general, ni el habitante del cuarto en particular, entienden ninguno de los signos chinos manipulados. Es más, para el habitante del cuarto, las figuras que manipula y despacha no son signos en absoluto pues no representan nada para él. El uso de reglas de manipulación sintáctica (computacional) no otorga al habitante del cuarto ninguna capacidad de interpretación semántica de los signos, ¿en qué sentido, entonces, dice Searle, se puede decir que el sistema así creado es un caso de competencia lingüística? Si descubriéramos que un dispositivo similar opera dentro de una persona a la que le adjudicamos comprensión de un lenguaje, nos veríamos obligados a dejar de decir que esa persona comprende el lenguaje utilizado. Por lo tanto, la afirmación de que la descripción funcional de un sistema basta para explicar y definir sus EM es falsa, pues por el experimento anterior se demuestra que el funcionalismo computacional resulta ciego a una característica que parece ser propia, y muy evidente, de algunos EM, a saber, su intencionalidad.

Hasta el momento, he expuesto tres argumentos (HEI, HQA y el experimento del cuarto chino) que pretenden demostrar la falta de identidad entre rol funcional y EM. La pretensión de quienes proponen estas objeciones es mostrar que el funcionalismo no es un programa teórico viable para explicar la mente. Creo, sin embargo, que ninguno de los tres problemas representa un argumento concluyente para lograr tal objetivo. Se hace necesario aquí dar una serie de opiniones respecto a los problemas planteados a los funcionalismos de segundo orden como el de Putnam.



En primer lugar hay que aclarar que el tema del contenido fenoménico representa un problema no sólo para los funcionalismos de segundo orden, sino también para el fisicismo-tipo y el conductismo lógico. En ninguna de estas teorías queda incluido el contenido cualitativo de los EM dentro de su definición. Por lo tanto, es posible construir contraejemplos similares con base en la existencia de los qualia que evidencien la incompletud en las definiciones respectivas. En segundo lugar, hay que recordar que los críticos del funcionalismo de segundo orden pretendían demostrar que no es válida la identidad entre EM y rol funcional y para ello se valen de la existencia del contenido cualitativo. No obstante, esta postura asume no sólo que tal contenido existe, sino que resulta un elemento necesario dentro de los EM. De no ser asumida esta afirmación, las críticas carecerían de la solidez buscada. Ahora bien, que efectivamente los qualia sean un elemento necesario dentro de los EM es debatible. Del hecho de que tal contenido exista y de que nos parezca en el análisis subjetivo que es el rasgo esencial de los EM no se sigue que efectivamente sea así. Además, otra vía es posible para el funcionalismo: aceptemos que el contenido cualitativo es consustancial a los EM, lo que habrá que hacer es buscar la forma de demostrar que la definición funcional sí es sensible a su existencia.<sup>33</sup> Una última posibilidad sería demostrar que no todos los EM poseen qualia. De tal modo, aun en caso de que la crítica demostrara que la definición funcional no logra capturar el

---

<sup>33</sup> Intentos se han hecho en ambos sentidos; para una defensa del funcionalismo por la vía de no conceder importancia al contenido cualitativo dentro de la definición funcional *vid*, P. y P. S. Churchland, "Function, qualia and intentionality" (*vid infra* p. 59); para un intento por demostrar que la definición funcional sí puede integrar el contenido cualitativo *vid* S. Shoemaker, "Funcionalismo y qualia".

contenido cualitativo de los EM esto no implicaría el completo fracaso del programa funcionalista.

Block y Fodor pretendían que las dos hipótesis esgrimidas en su texto demostrarían que “los estados psicológicos no pueden definirse *en absoluto* en términos funcionalistas”. Al parecer, esto ha de entenderse como que sería imposible definir funcionalmente cualquier EM. Creo, dado lo anterior que, la HQA y la HEI no bastan para lograr demostrar lo que los autores se proponían. Se ha mostrado cómo las dos objeciones en cuestión serían un problema para la definición funcionalista de los estados cualitativos, pero o se asume sin demostración previa que todos los EM poseen tal contenido o la crítica no prueba que ningún EM sea definible funcionalmente. No obstante, tampoco creo que el tema de los qualia pueda ser resuelto desde una perspectiva funcionalista por más que se haga una presentación reformulada de la teoría. En lo particular, creo que el rol funcional es un elemento importante a ser considerado dentro de la definición de los EM, pero que, por sí solo no puede capturar el contenido cualitativo de la experiencia debido a su énfasis en las relaciones abstractas de los EM y de sus entradas y salidas.

En segundo lugar, hay que señalar que el tema de la intencionalidad no presenta el mismo tipo de problemas al funcionalismo que el contenido cualitativo. Creo que una teoría funcionalista de segundo orden puede llegar a dar una explicación coherente de la intencionalidad de los EM. De hecho, hay versiones de la teoría que afrontan este problema con distintos resultados.<sup>34</sup> Lo cierto, es que tales

---

<sup>34</sup> Vid, M. García Carpintero, *op cit.*

versiones parecen comprometerse con una teoría representacional de la mente que, por sí misma, presenta múltiples problemas.<sup>35</sup>

Examinaremos a continuación una propuesta que parte desde otra perspectiva, pero que aún conserva la estructura reductora que presenta el funcionalismo de segundo orden; se trata en este caso de una reducción de los EM a la adaptación biológica. Trataré de mostrar las razones por las que tal teoría naturalizada de Shapiro no logra tampoco el objetivo buscado de definir por completo los EM.

---

<sup>35</sup> Vid, J. Fodor, *Concepts. Where cognitive science went wrong*; H. Putnam, *Realidad y representación*.

## 5. UN NUEVO INTENTO: “LA MENTE COMO ADAPTACIÓN”

Las críticas del funcionalismo hacia el fisicismo-tipo y al conductismo parecen mostrar las inconsistencias de estas propuestas. A su vez, los contraejemplos de los qualia ausentes, el espectro invertido y el cuarto chino evidencian consecuencias indeseables del propio funcionalismo. Sin embargo, aunque las tres teorías puedan ser desacreditadas como explicaciones completas de los EM, no podemos negar que parece subsistir alguna validez en las intuiciones que les dieron origen, ¿hay alguna forma de construir una teoría que conserve estas intuiciones justificándolas sobre terreno firme, pero que a su vez logre evitar los defectos que dieron lugar a los contraejemplos expuestos?

Lawrence A. Shapiro piensa que sí es posible construir tal teoría; él la llama: *La mente-como-adaptación*.<sup>1</sup> Según Shapiro, tanto los problemas que presentan el fisicismo como el funcionalismo, así como las ideas que los inspiran son sólo *intuiciones*. Las dos intuiciones que hacen problemática la aceptación del fisicismo y del funcionalismo se pueden exponer así:

- a. *Contra el fisicismo-tipo*: parece que es lógicamente posible que distintas instancias físicas presenten mente (argumento de la múltiple realizabilidad).

---

<sup>1</sup> A. L. Shapiro, “Mind the adaptation”, *Naturalism, evolution and mind*, pp. 23-41.

- b. *Contra el funcionalismo*: parece que no cualquier sistema que presente un rol funcional es una instancia de mente “Las mentes deben ser más que *únicamente* una relación abstracta entre inputs, otros estados y outputs.”<sup>2</sup>

Por otra parte, querríamos que nuestra teoría rescatara también este otro par de intuiciones:

- c. *del fisicismo*: la instancia física importa y, por lo tanto, ha de ser considerada en la definición de los EM.
- d. *del funcionalismo*: el rol funcional es necesario en la definición de los EM.

El conjunto de intuiciones a)-d) no es inconsistente, sostiene Shapiro, y puede ser estructurado y fundamentado en una teoría de los EM si nos servimos de la ayuda de algunos conceptos y técnicas de la biología y si estamos dispuestos a redefinir algunas de nuestras concepciones sobre la mente, “por esta razón, debemos lanzarnos ávidamente sobre un marco que prometa justificar nuestras intuiciones sin que él mismo apele (directamente, al menos) a otras intuiciones.”<sup>3</sup>

El método al que apela Shapiro es la cladística. El objetivo de este método de clasificación, dice, “es capturar las relaciones filogenéticas entre especies.”<sup>4</sup> La cladística, como parte de la sistemática, está comprometida con cierta forma de clasificación de los organismos; el cladista se basa en la evidencia de caracteres comunes entre linajes para hacer juicios sobre la existencia de un ancestro común más

---

<sup>2</sup> *Ibid.*, p. 24 (énfasis mío).

<sup>3</sup> *Ibid.*, p. 23.

<sup>4</sup> *Ibid.*, p. 25.

cercano o lejano entre dos linajes respecto de un tercero. La cladística parte de la suposición de la heredabilidad de los caracteres, de modo que, salvo mutación o presión selectiva, los descendientes de un ancestro común presentarán los mismos caracteres de éste. Ahora bien, es posible que, gracias a las presiones selectivas, dos individuos posean el mismo carácter sin haberlo heredado de un ancestro común. Este es el problema cladístico de la distinción entre homologías (cuando el carácter es heredado de un ancestro común) y homoplasias (cuando el carácter se presenta en ambos casos por evolución convergente, es decir, cuando el carácter no está presente en ambos linajes en virtud de su descendencia de un ancestro común, sino gracias a que las presiones evolutivas actuaron de forma semejante en ambos casos).

Existe, además, el problema de saber qué caracteres ha de considerar el cladista, de entre todos los que posee un organismo, para establecer relaciones filogenéticas. Hay la posibilidad de que un carácter presente en un individuo no le haya reportado ventaja evolutiva alguna a su especie en su historia filogenética, pero que, al tampoco haber representado desventaja alguna, se continuara transmitiendo a los descendientes. Para hacer la distinción entre ambos tipos de caracteres (el seleccionado evolutivamente por su utilidad y el heredado por su neutralidad) Shapiro hace uso de dos conceptos de Elliot Sober:

- a) *Selección de:* caracteres que se encuentran presentes al no haber sido eliminados por las presiones selectivas, pero que no están presentes *en virtud* de su utilidad.

- b) *Selección por*: aquellos caracteres que fueron seleccionados por la presión selectiva del medio por reportar utilidad evolutiva a la especie, i. e., mayor *adecuación*.<sup>5</sup>

Son los caracteres que cubren el requisito b) los que importan al cladista y los que, dice Shapiro, cuentan como adaptaciones. Aquí Shapiro introduce una definición del concepto de adaptación que será esencial para su propuesta. Para él, *adaptación* es el compuesto de una estructura más una función en virtud de la cual fue seleccionada evolutivamente tal estructura. “Usaré el término *función* para nombrar el efecto de una estructura que explica, si la estructura es una adaptación, por qué fue seleccionado ese tipo de estructura.”<sup>6</sup> De tal modo, el cladista, en la interpretación de Shapiro, puede hacer juicios de comparación (de semejanza o diferencia entre dos individuos respecto de un tercero) con base en dos ejes: estructura y función. Hay entonces una matriz de cuatro posibilidades: 1. Función y estructura diferentes; 2. Estructura diferente y función similar; 3. Estructura similar y función diferente y 4. Estructura y función iguales.

Este es el marco teórico de Shapiro desde la biología evolutiva. La afirmación central ahora es: *la mente es una adaptación*, i. e., la mente es una estructura más una función en virtud de la cual fue elegida evolutivamente tal estructura. En esto radica

---

<sup>5</sup> El término “adecuación” traduce el inglés “fitness”. *Fitness* se entiende como una medida para indicar los efectos de las presiones selectivas en el organismo; el término “adecuación” se usa aquí en este sentido técnico. Tal capacidad de adecuación se define como: a) la habilidad del organismo para llegar a la edad adulta (edad fértil) y b) la capacidad reproductiva del organismo. Para el concepto de *fitness* vid. E. Sober, *Philosophy of biology. Dimensions of philosophy*, cap. 3.

<sup>6</sup> *Ibid.*, p. 28.

el núcleo básico de la teoría en cuestión, ¿cómo logra esta teoría satisfacer el conjunto inicial de intuiciones a) - d)? Veamos.

El inciso a) era la intuición del funcionalismo contra el fisicismo-tipo que intentaba demostrar que los EM no eran idénticos con una estructura cerebral determinada. Shapiro intenta demostrar lo mismo, pero con un argumento complementario: dos estructuras iguales (estructuras cerebrales) pueden diferir en sus historias evolutivas y, por tanto, tener diferente función (rol funcional). Cada argumento, por sí solo, es suficiente para demostrar la no equivalencia entre estructura y función necesaria para el fisicismo-tipo; la diferencia, dice Shapiro, es que el primero se basa sólo en una intuición, mientras que el segundo está respaldado por el método cladístico.

Por su parte, las intuiciones c) y d) quedan justificadas por la definición misma de adaptación de Shapiro: si la adaptación es una estructura más una función, i. e., si función y estructura son en conjunto elementos necesarios y suficientes para que haya adaptación, entonces se infiere que tanto la instancia física como el rol funcional son importantes y han de ser considerados dentro de la definición de los EM. Afirmar que sí importa la instancia física dentro de la definición de los EM es acotar de facto la tesis de la múltiple realizabilidad: sí es posible que otras *cosas* tengan mente, siempre y cuando sean *cosas* no muy distintas a nosotros, de hecho, siempre y cuando sean bastante semejantes, es decir, siempre que tengan una historia evolutiva muy similar: “Así, cuando son concebidas como una adaptación, las mentes que tienen estructuras que difieren de la nuestra no son mentes. Nada salvo un cerebro puede ser una mente.



Es por esta razón que la nación de China o la economía de Bolivia no pueden ser mentes”<sup>7</sup>

Con esto, queda también satisfecho el inciso b). Shapiro busca la justificación de este nuevo chauvinismo de lo mental en la teoría de la evolución: la naturaleza tiene un conjunto pequeño de estructuras posibles para lograr las mismas funciones. Hablando de la múltiple emergencia de los ojos como órgano de la visión, Shapiro dice: “¿Por qué debemos encontrar en la naturaleza el mismo diseño repetido una y otra vez? Una respuesta plausible a esta pregunta es que sólo hay un puñado de estructuras que pueden funcionar como un ojo.”<sup>8</sup>

Lo anterior parece llevar a la conclusión de que “la función restringe la estructura”, por lo cual, la múltiple realizabilidad irrestricta de los EM, afirma Shapiro, es falsa. Si computadoras o extraterrestres se comportan como nosotros, entonces es altamente probable (debido a la complejidad que presenta esa adaptación que conocemos como “mente”) que computadoras o extraterrestres tengan nuestra misma estructura. “Si concedemos, como una extensión razonable a la generalización previa de que entre más compleja es la función, en mayor medida ésta restringe la estructura, entonces debemos ver como *muy* probable que extraterrestres y sistemas artificiales que presenten mentalidad tengan estructuras como-cerebros dentro de sus cabezas”<sup>9</sup> La mente, dijimos, es una estructura más una función, pero resulta que no es cualquier estructura, entonces sería más adecuado decir: la mente es *cerebro* más una función.

---

<sup>7</sup> *Ibid.*, p. 33.

<sup>8</sup> *Ibid.*, p. 34.

<sup>9</sup> *Ibid.*, p. 35.

## 6. LOS PROBLEMAS DE “LA MENTE COMO ADAPTACIÓN”

¿Logra Shapiro su objetivo?, ¿es, efectivamente, la teoría de la mente como adaptación el punto intermedio entre el fisicismo-tipo y el funcionalismo que permite rescatar las intuiciones que deseamos conservar y rechazar aquellas que consideramos como injustificadas de tales teorías?, ¿es correcto el uso que hace Shapiro de la biología evolutiva y del método cladístico para construir el marco teórico de su propuesta? A continuación presento una lista de seis objeciones a la teoría de Shapiro por las que creo que la respuesta a las preguntas anteriores es negativa en todos los casos. Tales objeciones se pueden dividir en dos grupos: aquellas que hablan sobre la incomprensión e incorrecto uso de los términos y los métodos de la biología evolutiva y aquella (la última) que tiene que ver con la pretensión reduccionista de la teoría de Shapiro.

### 1. *Incorrecta concepción de la cladística*

Como vimos, el uso del método cladístico es esencial para la propuesta de Shapiro, sin embargo, su comprensión de tal método parece no ser del todo adecuada. Según Shapiro, la cladística tiene como objeto capturar las relaciones filogenéticas entre *especies*. Pero la cladística no habla de especies. No es propiamente con especies con lo que trabaja tal método, sino con *grupos monofiléticos*. Un grupo monofilético es aquél que está formado por todos los descendientes de un ancestro común y sólo por estos.<sup>10</sup> Dado que al inicio de un grupo monofilético se encuentra un grupo ancestral

---

<sup>10</sup> Un grupo parafilético, por el contrario, es aquel que excluye a algunos de los descendientes del ancestro común. Algunos otros métodos de la sistemática biológica distintos al cladístico se distinguen

que se bifurca en linajes diferentes, entonces no es válido igualar el concepto de especie con el de grupo monofilético, pues este último puede contener, de hecho, varias especies. Además, se pueden hacer tantos grupos monofiléticos dentro del cladograma como lo decida el investigador; se puede crear un grupo monofilético siempre que haya un punto del esquema en donde se bifurcan los linajes. En resumen, la cladística y sus objetivos se definen sin apelar en absoluto al esquivo concepto de especie: “La sistemática filogenética, o cladística, estudia la diversidad orgánica a través del reconocimiento de las relaciones genealógicas de los organismos, los que se reflejan en la clasificación natural de los mismos.” Y sus objetivos:

1. Proveer, mediante la clasificación, el marco conceptual mediante el cual los biólogos pueden comunicar información acerca de los seres vivos.
2. Proporcionar, mediante los cladogramas, las bases para proveer diferentes interpretaciones evolutivas.
3. Predecir, mediante los cladogramas y las clasificaciones derivadas de los mismos, propiedades de los organismos recién descubiertos o poco conocidos.<sup>11</sup>

Una vez que el cladista establece el grupo monofilético con el cual trabajará, es necesario que defina los caracteres que le servirán como evidencia para establecer las relaciones genealógicas dentro del cladograma. En este punto surge un problema arduo en la filosofía de la biología y que merece ser tratado por separado.

## 2. ¿Qué califica como carácter?

En la biología evolutiva resulta sumamente vago el concepto de carácter “Un carácter [dice Murrone] es *cualquier* atributo (morfológico, celular, etológico, ecológico, biogeográfico, etc.) que podamos observar en un organismo.”<sup>12</sup> Bajo esta vaguedad igual califica como carácter la talla, la forma o el color del organismo, como toda la

---

de éste último por la aceptación de grupos parafiléticos. Vid J. Murrone, *El lenguaje de la cladística*, p. 11; E. Sober, *Reconstructing the past*, cap. 1.

<sup>11</sup> J. Murrone, *op. cit.*, p. 19.

<sup>12</sup> *Ibid.*, p. 19. Énfasis mío.

amplia gama de conductas de apareamiento o de reacción ante el peligro, así como órganos o miembros tales como los ojos o las alas. Quienes hacen psicología evolutiva tienden a valerse de esta indefinición del término para postular que los EM o, incluso, la mente, son caracteres heredables. ¿Resulta válida tal extensión del concepto? La cladística es una forma de clasificar organismos de acuerdo a sus relaciones filogenéticas, i. e., históricas. Detrás del método cladístico está la presunción evolutiva de la heredabilidad de los caracteres. De acuerdo con esto, podemos afirmar que un carácter que resulte de pertinencia para el estudio filogenético, ha de tener, cuando menos, las siguientes dos características: a) tener una codificación genotípica (que, eventualmente, puede traducirse en una expresión fenotípica)<sup>13</sup> y b) ser ventajoso o, cuando menos, neutral para la adecuación del organismo que lo posee. ¿La mente (o los EM, en todo caso) puede ser un carácter heredable en este sentido?, ¿qué podría significar tal afirmación?, ¿es posible, acaso, encontrar la codificación genotípica de la mente o de cada uno de los EM de un organismo? La respuesta a estas preguntas no se encuentra en la exposición de Shapiro y parecería que habría problemas al intentar aclarar estos puntos. Lo cierto es que, a pesar de la indeterminación en que parece encontrarse el concepto de carácter dentro de la literatura relevante, sí parece haber un gran salto conceptual al afirmar que “la mente” (o los distintos EM) es un carácter evolutivamente heredable en el mismo sentido que el color o la talla. Trataremos más este problema en la siguiente objeción.

---

<sup>13</sup> Cabe aclarar que no todo estudio dentro de la sistemática evolucionista se ha realizado atendiendo al requerimiento de este inciso. Es práctica reciente el estudio de las relaciones filogenéticas desde la genética molecular. Tal uso de la genética permite garantizar en mayor medida el requisito de heredabilidad (a mayor semejanza genética entre dos linajes, mayor probabilidad hay de herencia de los caracteres). De modo que, aunque es posible llevar a cabo un estudio filogenético sin tal requisito, no se garantiza de tal modo la exclusión de hipótesis de convergencia entre caracteres.

### 3. ¿Es la mente una adaptación?

¿Qué es lo que quiere decir exactamente que la mente sea una adaptación? Se dice de un carácter que es una adaptación, cuando la selección natural explica su presencia de una generación a otra.<sup>14</sup> Un carácter adaptado reporta mayor adecuación<sup>15</sup> al organismo que lo posee. Shapiro define las adaptaciones en términos de una estructura y una función, en donde ésta última es la razón por la cual fue seleccionada la estructura. Un primer problema es que en la exposición de Shapiro no queda claro qué es exactamente lo que se selecciona: la estructura, la función o el conjunto de ambas. Al inicio, parece que es la estructura la seleccionada en virtud de la función, sin embargo, luego se nos dice que dos estructuras iguales con diferente función son distintas adaptaciones, por lo que pareciera que el elemento importante es la función. Finalmente, con la afirmación doble de que la función restringe la estructura y la estructura, a su vez, restringe la función, parece que lo seleccionado es el compuesto de ambos. El problema es que la mente, tal como la entiende Shapiro, cae en el último de estos casos: es el compuesto de estructura y función. No hay mente si alguno de los términos difiere de la forma que adoptan en nuestra propia especie. Veamos por qué resulta problemática esta afirmación. El problema es relevante porque, de la manera en que se presenta el concepto de adaptación, no se deja lugar para conceptos como los de *neutralidad del carácter* o el de *exaptación*<sup>16</sup> que resultan de utilidad en la biología evolutiva.

---

<sup>14</sup> Vid., K. Sterelny y P. Griffiths, *Sex and Death*, p. 217.

<sup>15</sup> Vid., supra n. 31.

<sup>16</sup> Es neutro un carácter que si bien no reportó utilidad alguna en la adecuación del organismo, tampoco resultó desventajoso. Por lo tanto, la selección natural no lo “eligió”, pero tampoco lo “eliminó”. Por otra parte, “exaptación” es un término introducido por Vrba y Gould (*Vid.*, S. J. Gould, “Darwinism and the expansion of the evolutionary theory”, p. 383.) para designar una estructura cuya función actual no es

Hay dos posibilidades al decir que la mente es una adaptación: a) es la mente, como un todo, lo que reporta mayor adecuación o b) cada uno de los EM es una estructura con una función determinada que reportó evolutivamente mayores ventajas al individuo. Hacia el final del artículo, Shapiro afirma que lo que tiene en mente es b):

Por supuesto, las mentes no son cosas simples. 'Mente', de la manera en que uso el término, se refiere a una colección de capacidades cognitivas: memoria, atención, lenguaje, percepción, etc. Entonces, de forma más precisa, la sugerencia que estoy defendiendo es que *cada una de esas capacidades debe ser concebida como un tipo particular de estructura con una función particular*. [...] De acuerdo con esto, la 'mente-como-adaptación' es una abreviatura. Es un eslogan que nos invita a concebir las variadas capacidades cognitivas, que asociamos típicamente con la mente, como tipos de estructuras con funciones.<sup>17</sup>

La mente, entonces, es un conjunto de adaptaciones. ¿Cómo es esto posible? Si bien es cierto que hay algunas conductas que pueden ser indicios de ciertos EM (la conducta de huida ante la percepción de un depredador evidenciaría *miedo* y ciertas *creencias* y *deseos*, por ejemplo) que reportan mayor adecuación al organismo, también es cierto que hay EM para los que parecería absurdo hacer esta afirmación. Como afirma Lewontin, ¿qué ventaja evolutiva le reportó a un homínido ancestral poder construir oraciones subordinadas?, ¿qué ventaja en el éxtasis religioso?<sup>18</sup> La conexión entre la existencia de estos EM y la mejora de las capacidades de sobrevivencia y de reproducción del individuo no es en absoluto clara; es, cuando menos, dudable que tal

---

aquella para la que fue seleccionada. Tal es el caso, por ejemplo, de la existencia de las alas en las aves y su función para volar.

<sup>17</sup> L. Shapiro, *op. cit.* p. 36 (énfasis mío).

<sup>18</sup> R. Lewontin, "The evolution of cognition: Questions we will never answer", *An invitation to cognitive science: Thinking*, p. 107-132.

conexión exista y que, por lo tanto, la historia evolutiva del organismo sirva para definir (por sí sola, al menos) la existencia de los EM en general.

Shapiro podría contestar: no todos los EM reportan (o tienen por qué reportar) ventaja evolutiva alguna; algunos EM serán caracteres neutros y otros tantos serán exaptaciones. Sin embargo, si se trata de caracteres neutros, entonces no reportan ventaja alguna al organismo, i. e., no fueron seleccionados *en virtud* de la función que realizan y, por lo tanto, no son adaptaciones. Por otro lado, si tales EM son exaptaciones, entonces tenemos el problema de saber con qué función comparamos esa misma estructura presente en distintos organismos o en el mismo organismo a lo largo de su historia evolutiva. Estos problemas se presentan, según creo, debido a que la propuesta de Shapiro no toma en cuenta la posibilidad de que un carácter sea reclutado para diferentes funciones a lo largo de la historia evolutiva. Además, la identificación de cada EM con una adaptación (en el sentido usado por Shapiro) abre la puerta a la pregunta: ¿cuál es la estructura y la función específica de cada EM? Esto nos lleva a problemas ya conocidos que son el tema de la siguiente objeción.

#### 4. *De vuelta al fisicismo-tipo*

Otro problema complejo dentro de la biología evolutiva es el de poder identificar en forma unívoca la función de una estructura determinada. De las múltiples cosas que puede realizar una estructura hay que distinguir aquélla en virtud de la cual tal estructura fue seleccionada por reportar mayor adecuación al organismo. El problema es que esto no resulta siempre evidente. Se pueden establecer hipótesis evolutivas verosímiles que expliquen la heredabilidad de un carácter en virtud de alguna función

determinada, pero con frecuencia no hay evidencia concluyente para tales hipótesis. Se podría hablar de una infradeterminación de las historias evolutivas por la evidencia. Además, se vuelve a presentar aquí el tema de las exaptaciones: si resulta difícil saber cuál es la función actual en virtud de la cual fue seleccionada una estructura, más difícil aún saber cuál fue su función originaria.

En el caso de la mente el problema se complica más, pues no sólo queda por explicar cuál sea su función (o la función de cada uno de los EM), sino que tampoco queda claro cuál pueda ser su “estructura”. En su análisis sobre la posible existencia de extraterrestres y computadoras que exhiban mente, Shapiro deja ver que la estructura aludida es, por supuesto, el cerebro. Si ese es el caso, tenemos que vuelven a surgir problemas que nos resultan familiares: es difícilmente concebible que se pueda encontrar una estructura específica (y sólo una) para cada EM. ¡Ese era, precisamente, el problema del fisicismo-tipo que intentábamos evitar! Tomemos el cerebro de un afásico con regeneración de las capacidades lingüísticas dañadas (en donde la función es adquirida por una región cerebral que originalmente no desempeñaba tal tarea). Si lo comparamos con el cerebro de una persona no afásica, tenemos entonces una misma función y dos estructuras distintas. Supongamos ahora que un grupo de individuos encuentra una ventaja evolutiva en la adquisición de las habilidades lingüísticas por parte de la nueva región cerebral, existirían diferencias de estructura, pero no de función, lo que claramente cuenta como un ejemplo de diferentes adaptaciones en la matriz de criterios comparativos de Shapiro. La plasticidad cerebral hace pensar que estos cambios son posibles en múltiples regiones del cerebro y con múltiples funciones en un rango relativamente grande. Si afirmamos



que se trata en todos los casos de distintas adaptaciones, entonces se diluye el concepto de mente-como-adaptación (en donde se presupone que se trata de una sola adaptación) en: múltiples-adaptaciones-que-comparten-una-misma-función, y todas esas adaptaciones serían ejemplos de un mismo EM. El propio Shapiro podría contestar que, a pesar de todo, se sigue tratando de la misma estructura: el cerebro, sin embargo, eso va en contra de que cada EM es una estructura distinta con una función propia (recordemos que estructura y función individualizarían a cada EM).<sup>19</sup> Para el funcionalismo, el EM era sólo la función (el rol funcional) lo cual le permitía hacer válida la intuición de la múltiple realizabilidad de los EM. Esto está en consonancia con la posibilidad de realizar la misma función en distintas regiones cerebrales. No obstante, Shapiro no puede apelar al argumento de la múltiple realizabilidad, pues él lo ha restringido al grado de hacerlo casi irrelevante. Y esto nos lleva a la siguiente objeción.

##### 5. *Contra el argumento de la limitación de las estructuras*

La función, dice Shapiro, restringe la estructura. Esta era la fórmula para evitar caer en el liberalismo del funcionalismo. Pero resulta que esta misma fórmula, como mencionamos, parece regresarnos de lleno al chauvinismo fisicista. Shapiro afirma que los ojos son un buen ejemplo de esta restricción de la estructura por la función: sólo parece haber un *pool* de 8 a 10 estructuras posibles para cumplir la función de ver. Entonces, si extraterrestres y computadoras se comportan como nosotros, es porque han de poseer estructuras muy semejantes a las nuestras. Con esto, Shapiro parece

---

<sup>19</sup> Si dos estructuras cerebrales pueden realizar la misma función y si una misma estructura cerebral puede realizar diferentes funciones, entonces ¿cómo la pareja función-estructura individualiza cada EM?

negar casi por completo la existencia de las homoplasias. Las alas de los murciélagos, de las aves y de los insectos sirven para volar, pero, a pesar de semejanzas estructurales, y del hecho de que a todos esos miembros los llamamos “alas”, resulta que son bastante diferentes: difieren en el material del que están formados, en su estructura y en sus historias evolutivas. Sólo comparten la función, pero, incluso, algunos de esos miembros comparten tal función sólo en forma derivada en virtud de ser exaptaciones. Que tales miembros posean algunas semejanzas y compartan la función no dice nada sobre la imposibilidad de desarrollar estructuras distintas para cumplir esa misma función, sino que sólo muestra que organismos distintos se han visto obligados a usar recursos similares ante presiones evolutivas similares. Si tales recursos y presiones se modificaran, cabría pensar consecuentemente que así lo harían también las estructuras resultantes. Tal variación de recursos y condiciones es, precisamente, la que cabría esperar de algunos presuntos seres que vivieran fuera de nuestro medio (planeta) y que exhibieran mentalidad. ¿Por qué presuponer, entonces, que las estructuras que poseyeran habrían de ser *muy* similares a las nuestras? De hecho, ¿no es de sorprender que, bajo las mismas condiciones y con los mismos recursos, haya entre 8 y 10 posibles arreglos distintos para lograr la función de “ver”?

#### 6. *La-mente-como-adaptación es una reducción*

Shapiro dice expresamente que su teoría es una teoría de la identidad en el mismo sentido que lo son el fisicismo y el funcionalismo. Es decir, su intento es “reducir” la mentalidad a los términos de función y estructura de la biología evolutiva. Los problemas presentados por las teorías expuestas hasta aquí muestran que el elemento reductor generalmente resulta insuficiente para definir el campo semántico de lo

mental. Lo mismo sucede en el caso presente. Aun aceptando que Shapiro tenga razón en su interpretación de los conceptos de “función”, “estructura” y “adaptación” y que todos los EM no son sino adaptaciones, todavía quedaría por explicar cómo es que no todas las adaptaciones son EM, es decir, aún está pendiente especificar cuál es ese rasgo que presentan unas adaptaciones en virtud de las cuales las llamamos mentales y que las hace diferentes del resto. De la misma manera en que el funcionalismo presenta una ceguera ante los qualia y la intencionalidad, podríamos decir que esta teoría naturalista presenta una ceguera ante la distinción mental – no mental. La teoría de Shapiro se pondría en camino de resolver este problema si fuera capaz de demostrar, precisamente, que resuelve los problemas de los qualia y de la intencionalidad, pues, a fin de cuentas, al rescatar los postulados funcionalistas y “corregirlos” debería mostrar cómo lidia con estos problemas heredados. Sin embargo, no parece haber pistas que indiquen que la-mente-como-adaptación es una teoría apta para afrontar estos retos.

En resumen, la teoría de Shapiro fracasa por dos razones: 1) la aplicación de los conceptos y métodos de la biología evolucionista resulta incorrecta lo cual provoca que varias de sus afirmaciones y de los argumentos que en ella se basan resulten inválidos o confusos (tal es el caso de su versión del argumento de la múltiple realizabilidad); 2) Shapiro intenta decirnos que los métodos cladísticos y los conceptos de la biología evolucionista bastan para explicar los fenómenos mentales; dada la afirmación de que esta teoría es también una “teoría de la identidad” hemos de suponer que para Shapiro hay una reducción ontológica y metodológica de por medio: “la mente no es más que un conjunto de adaptaciones” y “la mente se entiende

completamente a partir de los métodos de la biología evolucionista”, pero tal marco teórico deja sin resolver los problemas de la experiencia fenoménica y de la intencionalidad que ya presentaban el funcionalismo; además, no existe en la propuesta una pauta para explicar que si bien todas las cosas que exhiben mente son producto de adaptaciones, no todas las adaptaciones son instancias mentales. La adaptación por sí sola resulta insuficiente para explicar los EM y los métodos de la biología evolucionista (cuando menos en la forma usada por Shapiro) resultan incompletos para explicar el campo de lo mental.

## 7. ¿CUÁL ES EL PROBLEMA CON LAS TEORÍAS DE LA REDUCCIÓN?

En el presente trabajo he utilizado el término “reducción” para referirme a las teorías aquí expuestas. He preferido este término y no el de “identidad” por dos razones. En primer lugar, no todas las teorías que intentan llevar a cabo alguna forma de reducción son teorías de la identidad. El término “identidad” se identifica, como lo vimos, con un tipo específico de teorías, ya sean teorías fisicistas o funcionalismos de primer orden.<sup>20</sup> Los compromisos de las teorías de la identidad no necesariamente son compartidos por todas las teorías que aquí hemos revisado. De utilizar tal término, el equívoco sería difícil de eludir.

En segundo lugar, creo que hay una razón que pasa por el significado de ambos términos que hace necesario distinguirlos. Cuando decimos que A es idéntico a B queremos decir que todos los casos de A son casos de B y que todos los casos de B son casos de A. Para postular identidad en este sentido es necesario afirmar estas dos relaciones. Una teoría de la identidad debe de afirmar ambas relaciones. Si sustituimos A por EM, entonces sólo quedará por sustituir B con aquello que creamos que es efectivamente A, y entonces tendríamos una teoría de la identidad. Puesto de este modo, se muestra que la afirmación de las teorías de la identidad es muy fuerte, se trata de la afirmación de una relación simétrica: B ha de ser suficiente y necesaria para explicar A y viceversa. Típicamente, las teorías de la identidad son aquellas que llevan a cabo una reducción de carácter ontológico, es decir, intentan decirnos que lo que son

---

<sup>20</sup> Vid. C. D. Broad, *The Mind and its Place in Nature*; J. J. C. Smart, “The Identity Theory of Mind”; D. Lewis, “Un argumento en favor de la teoría de la identidad”.

ciertos términos (los mentales, por ejemplo) en realidad no es más que ciertos otros términos ya previamente conocidos de un campo teórico distinto (estados neurofisiológicos, por ejemplo). En este sentido, como mencionábamos arriba, las teorías de la identidad toman su modelo de las reducciones ontológicas de la ciencia. Son estas las teorías que hemos clasificado como teorías de primer orden.

Sin embargo, no todas las teorías de la reducción requieren afirmar tal simetría. Como se vio en el capítulo 3, hay otro tipo de teorías, las de segundo orden, que no intentan identificar los EM con estados físicos. Son las teorías que postulan el argumento de la múltiple realizabilidad. Para estas teorías no hay identidad entre EM y estados físicos, sin embargo, sigue existiendo un postulado reduccionista dentro de ellas. Decir que hay identidad entre EM y estados físicos podría ser una vía para decir qué son los primeros, pero al negar tal identidad queda por explicar entonces qué sean los EM. Al llevar a cabo esta explicación es que las teorías de segundo orden afirman una forma de reducción. Quizá no se trata de una reducción ontológica en el sentido de la identidad, pero sí de una reducción semántica, epistémica o metodológica. En el caso del funcionalismo computacional, la afirmación es: los EM son roles funcionales. García Carpintero deja en claro de qué tipo de reducción se trata cuando afirma: “No se trata de pretender que el concepto habitual de lo mental sea funcional, sino más bien, de proponer que se entienda así”.<sup>21</sup> Se trata entonces de reducir un campo epistémico (lo mental) a otro (el rol funcional). Y es a este nivel en donde intentan actuar las críticas de los qualia y de la intencionalidad. En este sentido, hay que decir

---

<sup>21</sup> M. García Carpintero, *op cit.* p. 58.

que tanto los funcionalismos computacionales como la tesis de Shapiro son teorías de segundo orden.<sup>22</sup>

Ahora bien, como mencionamos anteriormente, hay que tener en claro qué es exactamente lo que demostrarían las críticas a las teorías de segundo orden en caso de ser ciertas. A diferencia de las teorías de la identidad, las teorías de segundo orden no establecen una relación simétrica, por lo que, en principio, nadie exige que todos los roles funcionales sean EM o que todas las adaptaciones sean EM. El problema es que las críticas centran la atención en un problema que se presenta en este tipo de reducción: ¿cuál es entonces el criterio para saber cuándo un caso de B (el rol funcional o una adaptación) no debe ser entendida como un caso de A (un EM)?

Lo que aquí se afirma respecto a las objeciones a las teorías de segundo orden es lo siguiente, tales objeciones en caso de resultar ciertas, sólo indican que las teorías de segundo orden carecen del criterio adecuado para indicar cuándo la reducción es exitosa y, aun en caso de demostrar que tal criterio es imposible de encontrar dentro de la teoría en cuestión, sólo se demostraría que la definición usada es incompleta, pero no innecesaria. Creo efectivamente que el rol funcional y el elemento adaptativo habrían de resultar de importancia dentro de una definición completa de los EM, pero creo también que ni el funcionalismo computacional ni la teoría de la-mente-como-adaptación logran establecer por completo la necesidad de estos elementos dentro de

---

<sup>22</sup> Incluyo a la propuesta de Shapiro entre las de segundo orden pues tiene la intención expresa de dar un argumento científico para demostrar que es cierta la múltiple realizabilidad de los EM. Esto es así puesto que él considera que los argumentos dados hasta el momento son sólo “intuiciones”. Sin embargo, como se vio en el capítulo anterior su intento resulta fallido pues se queda a medio camino entre la múltiple realizabilidad y el fisicismo-tipo. Esto lo acercaría a una teoría de primer orden, pero, de cualquier modo, la reducción que intenta llevar a cabo no es una del tipo de la identidad.

tal definición por encontrarse ocupados en la búsqueda por un argumento que demuestre que no sólo son necesarios, sino suficientes para caracterizar los EM. Una vez que se abandone tal pretensión de suficiencia podemos dejar de hablar de definiciones funcionalistas o naturalistas y podremos comenzar a hablar de aportaciones funcionalistas o adaptativas con miras a una definición completa de los EM. Con estas ideas en mente, podemos finalmente plantear algunas cuestiones relacionadas con el tema de la reducción en la filosofía de la mente.



## 8. LA REDUCCIÓN EN LA FILOSOFÍA DE LA MENTE

Apelando a su éxito en el proceder científico, el tema de la reducción ha propiciado un amplio debate en la filosofía de la ciencia del siglo XX<sup>23</sup>. La filosofía de la mente, como hemos visto, no se ha encontrado ajena a este interés. Una de las formas de abordar el problema mente-cuerpo es afirmando que uno de los elementos de la relación se reduce al segundo (el cerebro) o a algún otro (la conducta, el rol funcional, etc.). En sus diferentes versiones, esta es la idea que subyace a las teorías que han sido revisadas en el presente trabajo. Sin embargo, “reducción” está lejos de ser un término de uso unívoco, de modo que, como bien señala Searle<sup>24</sup>, hemos de concluir haciendo algunas especificaciones en torno al uso de la reducción en la filosofía de la mente.

**¿Qué se entiende por reducción?** En la discusión tradicional en el ámbito de la filosofía de la ciencia, la reducción se ha entendido como “la afirmación de que un rango de fenómenos puede ser *asimilado* a algún otro rango de fenómenos aparentemente distinto.”<sup>25</sup> El problema es que no es claro qué se ha de entender por “asimilar” un rango de fenómenos a otro. Tal asimilación se ha entendido como asimilación causal: se demuestra que B es causa de A; como asimilación legal: las leyes de A se pueden sustituir por las leyes de B; como asimilación lógica: los términos de A se pueden traducir a términos de B; como asimilación ontológica: las entidades de A *no son más que* entidades de las que se habla en B; como asimilación epistemológica:

---

<sup>23</sup> Hay una amplia bibliografía sobre el tema de la reducción en la ciencia, *vid*, por ejemplo, R. Carnap, “Logical foundations and the unity of science”, P. Oppenheim y H. Putnam, “The unity of science as a working hypothesis”, E. Nagel, “Issues in the logic of reduction”, P. Feyerabend, “How to be a good empiricist”; A. Garfinkel, “Reductionism”, J. Dupré, “Reductionism”, etc.

<sup>24</sup> J. Searle, *El redescubrimiento de la mente*, cap. 5.

<sup>25</sup> J. Dupré, *op. cit.*, p. 402 (énfasis mío).

conocer el campo de fenómenos de B es conocer el campo de fenómenos de A; o como asimilación metodológica: Los métodos y herramientas utilizados para estudiar B son los mismos necesarios para estudiar A. ¿Cuál de estos distintos tipos de reducción es el que está implicado en la filosofía de la mente? Creo que la respuesta depende de la teoría en cuestión. El tipo de reducción más radical y la que presuntamente se ha mostrado como efectiva en la ciencia física es la ontológica. Pero no es este el tipo de reducción que impera en todas las teorías de la mente. Creo que es válido decir que hay reducción ontológica en el fisicismo-tipo y, por lo tanto, en las teorías de primer orden como el funcionalismo analítico de Lewis. El conductismo lógico, por su parte, es claramente un caso de reducción lógica en donde se intenta redefinir términos mentales por medio de términos conductuales.

En el caso de los funcionalismos de segundo orden me parece que la reducción utilizada es de carácter epistémico, pues lo que se busca es entender a los EM como roles funcionales. En el caso de Shapiro, a pesar de que su forma de justificar el argumento de la múltiple realizabilidad lo acerca más al fisicismo-tipo, creo que su teoría podría ser considerada como un intento de reducción metodológica: es la cladística, el mismo método de la biología evolucionista, la que nos permite definir, comparar e individuar los EM.

En última instancia, creo que la reducción ontológica implica los demás tipos de reducción, en ese sentido, pienso que si se demostrara la verdad de la reducción ontológica de la mente a alguna otra instancia, ninguna teoría de la reducción en cualquiera de los demás sentidos del término sería incompatible con ella. Si

efectivamente la mente no es más que X, entonces también es cierto que las leyes que explican la mente se asimilan a las leyes de X y que nuestro conocimiento completo de X nos da el conocimiento completo de la mente, etc. No obstante, me parece que se ha desplazado el interés de hacer una reducción ontológica a intentar reducciones de otro tipo en la filosofía de la mente.

**¿Por qué deseamos reducir?** Creo que si los intentos de reducción en la filosofía de la ciencia y en la filosofía de la mente han sido tan populares se debe a tres razones principales: 1) algunos de los más importantes éxitos científicos (particularmente en la ciencia física) parecen ser ejemplos de reducción; 2) hay un tipo de argumento que podríamos llamar “de la composición” que dice: la realidad física está compuesta de micropartículas; tales micropartículas son estudiadas por la física, la cual ha desarrollado leyes que explican su comportamiento; por lo tanto, las estructuras y propiedades que surjan de la composición de tales partículas estarán explicadas por la mismas leyes que rigen el comportamiento de éstas<sup>26</sup>; 3) una razón de carácter heurístico: resulta más sencillo y económico explicar un rango de fenómenos por medio de unos pocos principios y elementos; si logramos llevar a cabo la reducción de un nivel teórico a otro, no hay necesidad de multiplicar las leyes ni las entidades en nuestro universo científico.

No obstante, **¿es útil reducir en la filosofía de la mente?** Resultan evidentes las ventajas de llevar a cabo una posible reducción en la filosofía de la mente; tendríamos entonces la posibilidad de explicar el campo de lo mental a partir de principios ya

---

<sup>26</sup> Vid. H., Putnam y P. Oppenheim, *op. cit.*; J. Dupré, *op. cit.*

conocidos sin necesidad de multiplicar las entidades ni de inaugurar un “lenguaje científico” distinto a los lenguajes ya empleados, utilizando así las mismas leyes de un nivel teórico inferior, ya sean leyes de la conducta, de la neurofisiología o de la biología evolutiva. El problema es que el programa reduccionista se enfrenta con múltiples problemas. Ninguna de las tres razones antes expuestas para llevar a cabo la reducción resulta concluyente: 1) la historia y la filosofía de la ciencia han puesto en entredicho que los casos tradicionalmente aludidos como éxitos científicos sean en efecto casos de reducción (cuando menos, de reducción ontológica), además, hay quienes afirman que, aunque la reducción en física fuera cierta y exitosa, no cabría aplicar tal concepto a las teorías de otras ciencias como la biología;<sup>27</sup> 2) la idea de que las leyes que explican un rango de fenómenos pueden explicar un rango de fenómenos de mayor complejidad se enfrenta con el obstáculo de la traducción teórica y de las llamadas “leyes puente”, la cuestión es: si en principio las leyes de dos niveles teóricos resultan distintas, ¿qué enunciados o conceptos hay que agregar para que se efectúe la reducción en forma completa y exitosa?;<sup>28</sup> 3) finalmente, la economía explicativa resulta de valor científico, sin embargo, ¿por qué hemos de privilegiar la economía explicativa cuando ésta nos lleva a definiciones a todas luces incompletas?

Específicamente, en el campo de la filosofía de la mente, hemos intentado mostrar que las definiciones que emanan de las teorías de la reducción de los EM que hemos revisado resultan incompletas debido a la pretensión de suficiencia del nivel teórico que reduce a lo mental. Otro tipo de argumento ha sido postulado por autores

---

<sup>27</sup> Vid P. Kitcher, “1953 and all that: A tale of two sciences”.

<sup>28</sup> Vid E. Nagel, *op. cit.*; P. Feyerabend, *op. cit.*

como T. Nagel y F. Jackson y ha sido retomado por J. Searle<sup>29</sup> para sostener la idea de la irreductibilidad de la conciencia; a este argumento le podríamos llamar de “la necesidad del elemento subjetivo” o, para abreviar, “de la subjetividad”. La idea es que, si es cierto que han existido reducciones en la ciencia, éstas han sido exitosas gracias a que eliminan un elemento que no resulta necesario en los fenómenos a explicar, a saber, el componente subjetivo (lo que aquí hemos llamado contenido cualitativo o qualia); sin embargo, llevar a cabo tal tipo de reducción en el campo de lo mental significa eliminar un componente que es claramente necesario: no hay dolor, percepción o emoción sin el componente subjetivo que está implicado en estos EM. La idea es que, si eliminamos los qualia de la explicación, ya no logramos definir el campo de lo mental. De ser cierto, el argumento de la subjetividad mostraría que las teorías de la reducción de los EM fracasan en su intento, todas ellas o no logran explicar la existencia del componente subjetivo o de plano lo eliminan de la definición.

A pesar de lo plausible que pueda sonar un argumento como el de la subjetividad, no todos están de acuerdo en que se debería rescatar el componente subjetivo en la definición de los EM. Paul Churchland, por ejemplo, postula un tipo de reducción con su materialismo eliminativo: en lugar de reducir el campo de lo mental al campo de lo físico, substituyámoslo por completo en favor de una *neurofilosofía*. En esta perspectiva, los términos mentales no son más que un reducto pre-científico que resultan del todo inútiles para una comprensión científica de la mente. Que el componente subjetivo *parezca* necesario no significa que efectivamente lo sea.<sup>30</sup> Por

---

<sup>29</sup> Vid T. Nagel, “Cómo es ser un murciélago”; F. Jackson, “Lo que María no sabía”; J. Searle, *op. cit.*

<sup>30</sup> Vid P. S. Churchland, *Neurophilosophy*; P. y P. S. Churchland, *op cit.*

otro lado, hay a quienes el argumento de la subjetividad los lleva a una conclusión pesimista sobre la posibilidad de resolver el problema mente-cuerpo por quedar demostrado que es imposible llevar a cabo un estudio científico de un ámbito (el de lo mental) que es por esencia no público ni cuantificable.<sup>31</sup>

En lo personal, tiendo a creer que es posible caminar entre ambos extremos para construir una teoría de la mente no reductora y para no caer en el pesimismo desconsolador sobre nuestro conocimiento de la mente. En primer lugar, creo que hay una razón detrás del argumento de la subjetividad que va más allá de la pura intuición y que justifica un tratamiento diferente de los casos de reducción en la ciencia física: la posición de la primera persona.<sup>32</sup> Cada quien es capaz de reconocer la existencia de sus propios EM y de reconocer la particular forma en que se presentan, de modo que, cuando adjudicamos EM a otra persona lo hacemos precisamente con referencia a nuestra propia experiencia con tales estados. De modo que el hecho de identificar un elemento necesario en nuestra experiencia no resulta igual a identificar erróneamente un elemento necesario en un objeto externo.<sup>33</sup> No queremos reducir en

---

<sup>31</sup> C. McGinn, "¿Podemos resolver el problema mente-cuerpo?".

<sup>32</sup> Aquí hay que hacer dos aclaraciones: 1) si bien es cierto que está a debate si todos los EM presentan qualia, me parece evidente que todos tienen un componente subjetivo cuando son concientes: reconocemos su existencia y sus características; 2) aun cuando se ponga en entredicho la autoridad de la primera persona (como sucede en varios argumentos epistemológicos) creo que eso no demuestra la nula relevancia del reporte de la experiencia por parte del sujeto, incluso el reporte erróneo del sujeto sobre una experiencia propia será un signo importante a tomar en cuenta en el análisis y definición de los EM.

<sup>33</sup> P. y P. S. Churchland afirman en "Function, qualia and intentionality" que la identificación de los qualia como necesarios en la definición de los EM es una identificación incorrecta del mismo tipo que identificar como necesarias a las rayas del tigre para que un animal pertenezca a tal especie. Creo, sin embargo, que si bien es cierto que resulta coherente preguntar por la posibilidad de un tigre sin rayas, una pregunta como "¿cómo sería un dolor sin la sensación de dolor?" carecería de sentido.

el campo de lo mental porque no queremos perder un elemento que, si bien no es suficiente, sí parece necesario para la definición de los EM.<sup>34</sup>

En segundo lugar, podemos preguntar ¿por qué no conceder fiabilidad a la experiencia subjetiva? La respuesta parece clara: porque no resulta científicamente manejable. Pero que una parte de la realidad sea “científicamente manejable” no es una cualidad intrínseca a tal fragmento de la realidad, es una consecuencia de nuestra eficacia metodológica. Si la vivencia subjetiva de los EM no resulta científica se debe a que no hemos logrado crear una metodología que supere la brecha explicativa haciendo científico lo subjetivo. Me parece que, frente a las complicaciones a las que se enfrentan las teorías de la reducción y ante el amplio campo de estudio que han abierto las llamadas ciencias cognitivas desde hace algunos años, resulta de especial importancia dedicar los esfuerzos a la creación de una metodología que logre meter de lleno a la experiencia subjetiva dentro del debate científico, reconociendo la necesidad de tal experiencia para la definición de los EM en el mismo sentido en que la biología evolutiva, la psicología, la inteligencia artificial y las neurociencias ponen el énfasis en la necesidad de integrar la historia evolutiva de la especie, la conducta, el rol funcional y la base neurológica dentro de tal definición.

Algunos intentos se han hecho ya en ese sentido. Tal es el caso de la *heterofenomenología* de Daniel Dennett<sup>35</sup> que ha suscitado una amplia discusión en los últimos años. Sin embargo, queda aún por estudiar si métodos como el referido

---

<sup>34</sup> Vid J. Searle, op. cit., p. 128 y ss.

<sup>35</sup> Vid D. Dennett, “La conciencia explicada”; *id.* “How to study human consciousness empirically, or nothing comes to mind”; *id.* “Who’s on first?, heterophenomenology explained”; *id.* “Heterophenomenology reconsidered”.

resultan útiles para dar un tratamiento científico a la experiencia subjetiva por medio de los reportes en primera persona una vez que son puestos en práctica. En cualquier caso, creo que, por lo pronto, la tarea actual de las ciencias cognitivas ha de enfocarse específicamente en la construcción de metodologías que logren establecer puentes, criterios de traducción y aplicación recíproca de conceptos y de leyes entre las muy diversas disciplinas que están implicadas en el debate. Si logramos cumplir tal propósito, podríamos vislumbrar un camino para escapar tanto de las restricciones del reduccionismo como del pesimismo de quien piensa al problema mente-cuerpo como un dilema irresoluble. En resumen, sobre este esfuerzo de las ciencias cognitivas podríamos repetir lo que Brentano afirmaba respecto de la psicología de su tiempo:

No hay rama de la ciencia que haya dado frutos más pobres para nuestro conocimiento de la naturaleza y la vida, y no hay ninguna otra de la que esperemos la satisfacción de nuestras necesidades más esenciales. No hay área del conocimiento, excepto únicamente la metafísica, que la gran masa de las personas mire con mayor desprecio, y ninguna a la que ciertos individuos atribuyan mayor valor y que tengan en más alta estima. Y aun más, el reino entero de la verdad parecería pobre y desdeñable para muchos si no se lo definiera de manera que incluya esta esfera del conocimiento y toda otra ciencia no le parece digna más que en la medida en que prepare el camino de esta. Y las otras ciencias son, de hecho, sólo la base; la psicología es como la coronación que concluye la obra. Todas las otras ciencias son una preparación para la psicología; ésta depende de todas pero, por su parte, es necesario que ejerza una influencia más poderosa sobre ellas. Es necesario que renueve la vida entera del hombre, que acelere y asegure el progreso. Y si, por un lado, parece ser la cumbre sobresaliente de la estructura de la ciencia, por otro, está destinada a volverse la base de sociedad y de sus posesiones más nobles, y, por esto mismo, a volverse también la base de todo el esfuerzo de los investigadores.<sup>36</sup>

Debido a la interacción de múltiples disciplinas, parece que el campo del estudio de la mente se encuentra confuso, y desorganizado, aunque con gran actividad. Hoy, a diferencia de la época en que hablaba Brentano, no es una disciplina la que participa en la investigación. Quizá nos encontramos en el mismo nivel pre-

---

<sup>36</sup> F. Brentano, *op. cit.* p. 15-16.



teórico que en aquél entonces, pero la diferencia es que, debido a lo complejo de la interacción entre las disciplinas, los retos metodológicos serán mucho mayores.

## BIBLIOGRAFÍA

- Armstrong, David, M., *A materialist theory of the mind*, London, Routledge, 1968.
- Block, Ned, 'Troubles with functionalism', en *Readings in the philosophy of psychology*, Vol. I, Harvard University Press, Cambridge, MA, 1980.
- , "Are absent qualia impossible?", en *The philosophical review*, vol. 89, no. 2, pp. 257-274, 1980.
- Block, Ned y Jerry Fodor, "Lo que no son los estados psicológicos", Maite Escurdia y Olbeth Hansberg (comps.), *La naturaleza de la experiencia. Vol. I Sensaciones*, trad. Ana Isabel Stellino, IIF-UNAM, México, 2003.
- Boller, Francois, "Rational basis of rehabilitation following cerebral lesions: A review of the Concept of cerebral plasticity", *Functional neurology*, Vol., 9, no. 2, pp. 65-72, 2004.
- Brentano, Franz, *Psicología desde un punto de vista empírico*, trad. Hernán Sholten, en internet (Fuente: Brentano, Franz: *Psychologie du point de vue empirique*, Paris, Éditions Mouton, 1944. Cotejado con la versión alemana: *Psychologie von Empirischen Standpunkt*, Leipzig, Felix Meiner Verlag, 1924.)
- Broad, C. D., *The mind and its place in nature*, London, Routledge and Kegan Paul, 1937.
- Buñuel, E. y E. Carbonell, "Conducta simbólica y neuroplasticidad: ¿Un ejemplo de coevolución gen-cultura?", *Revista de neurología*, vol.1, no. 39, pp. 48-55, 2004.
- Byrne, Alex y David R. Hilbert, "How do things look to the color blind?", *Color ontology and color science*, J. Cohen and M. Matthen (eds.), MIT Press.
- Carnap, Rudolph, "Logical foundations and the unity of science", R. Boyd, P. Gasper y J. D. Trout (eds.), *The philosophy of science*, 1991.
- Campbell, Keith, *Cuerpo y mente*, IIF-UNAM, México, 1988.
- Churchland, Paul, *Materia y conciencia. Introducción contemporánea a la filosofía de la mente*, trad. Margarita N. Mizraji, Gedisa Editorial, 1999.
- , *Neurophilosophy*, Bradford Books/MIT Press, 1986.
- Churchland, Paul y Patricia Churchland, "Function, qualia and intentionality", J. I. Biro y W. Shahan (eds.), *Mind, brain and function: essays in the philosophy of mind*, University of Oklahoma, 1982.
- Crane, Tim, "Intentionality as the mark of the mental", Anthony O'Hear (ed.), *Contemporary issues in the philosophy of mind*, 1998.

- Davidson, Donald, "Sucesos mentales", *Ensayos sobre acciones y sucesos*, trad. Olbeth Hansberg, José Antonio Robles, Margarita Valdés, IIF-UNAM, Crítica #26, 1995.
- Dennett, Daniel, *La conciencia explicada*,
- , "Heterophenomenology reconsidered", *Phenom Cogn Sci*, 2007.
- , "How to study human consciousness empirically, or nothing comes to mind", *Synthese*, 59, pp. 159-180, 1982.
- , "Quining Qualia", W. Lycan (ed.) *Mind and Cognition*, Oxford : Blackwell, 519-548, 1990.
- , "Who's on first?, heterophenomenology explained" en *Journal of Consciousness Studies*, 10, pp.19-30, 2003.
- Dupré, John, "Reductionism", W. H. Newton- Smith (ed.), *A companion to the philosophy of science*, Blackwell, 2001.
- Fitzhugh, Kirk, "Fact, theory, test and evolution", *Zoologica scripta*, vol. 37, no. 1, pp. 109-113, 2008.
- Feyerabend, Paul K., "How to be a good empiricist", M. Curd y J. A. Cover (eds.), *Philosophy of science. The central issues*, 1998.
- Fodor, Jerry, *Concepts. Where the cognitive science went wrong*, Clarendon Press, Oxford, 1988.
- García Carpintero, Manuel, "El funcionalismo", Fernando Broncano (ed.), *La mente humana*, Enciclopedia Iberoamericana de Filosofía, Trotta, Madrid, pp. 43-76, 1995.
- Garfinkel, Alan, "Reductionism", R. Boyd, P. Gasper y J. D. Trout (eds.), *The philosophy of science*, 1991.
- Gould, Stephen, J., "Darwinism and the expansion of evolutionary theory science", *Science*, New series, vol. 216, no. 4544, pp. 380-387, 1982.
- , "Introduction", T. Dobzhansky, *Genetics and the origin of species*, The Columbia classics in evolution series, pp. XVII-XLI, 1982.
- Graham, George, "Behaviorism", *Stanford encyclopedia of philosophy*, <http://plato.stanford.edu/entries/behaviorism/>, 2000.

- Griffiths, Paul. E., "Cladistic classification and functional explanation", *Philosophy of science*, vol. 61, no. 2., pp. 206-227, 1994.
- Guttenplan, Samuel (ed.), *A Companion to the philosophy of mind*, Blackwell, Oxford, 1994.
- Jackson, Frank, "Lo que María no sabía", Maite Escurdia y Olbeth Hansberg (eds.), *La naturaleza de la experiencia. Vol. I Sensaciones*, trad. Ana Isabel Stellino, IIF-UNAM, México, 2003.
- Jacquette, Dale, "Brentano's concept of intentionality", en *The Cambridge companion to Brentano*, Cambridge University Press, 2004.
- Kim, Jaegwon, *Mind and supervenience*, Cambridge University Press, Cambridge, 1993.
- , *Philosophy of mind*; Westview Press, 1996.
- Kitcher, Philip, "1953 and all that: A tale of two sciences", M. Curd y J. A. Cover (eds.), *Philosophy of science. The central issues*, 1998.
- Levin, Janet. "Functionalism", *Stanford encyclopedia of philosophy*, <http://plato.stanford.edu/entries/functionalist/>, 2004.
- Lewin, Roger, "Is your brain really necessary?", *Science*, vol. 210, no.12, pp. 1232-1234, 1980.
- Lewis, David, "How to define theoretical terms", *Philosophical papers*, vol. 1, OUP, Oxford, 1983.
- , "Mad pain and martial pain", *Philosophical papers*, vol 1, OUP, Oxford, 1983.
- , "Un argumento en favor de la teoría de la identidad", trad. Enrique Villanueva, Cuadernos de Crítica, IIF-UNAM, México, 1984.
- , "Psychophysical and theoretical identifications", *Australasian Journal of philosophy*, 50, pp. 249-258, 1972.
- Lewontin, Richard, "The evolution of cognition: Questions we will never answer", D.N. Oshoson y E.E. Smith (eds.), *An invitation to cognitive science: Thinking*, Vol 3. p. 107-132, MIT Press, Cambridge, pp. 107-132, 1990.
- Lycan, William, "The mind-body problem", Stephen Stich y Ted Warfield (eds.), *The Blackwell guide to philosophy of mind*, Blackwell, Oxford, Ch. 2, pp. 47-64, 2003.
- Mayr, Ernst, *One long argument*. Harvard University Press, Cambridge, MA, 1991.

- McGinn, Collin, “¿Podemos resolver el problema mente-cuerpo?”, Maite Ezcurdia y Olbeth Hansberg (comps.) *La naturaleza de la experiencia. Vol. I Sensaciones*, trad. Héctor Islas, IIF-UNAM, México, 2003.
- Morrone, Juan, *El lenguaje de la cladística*, UNAM, México, 2003.
- Nagel, Ernst, “Issues in the logic of reduction”, M. Curd y J. A. Cover (eds.), *Philosophy of science. The central issues*, 1998.
- Nagel, Thomas, “¿Cómo es ser un murciélago?”, Maite Ezcurdia y Olbeth Hansberg (comps.) *La naturaleza de la experiencia. Vol. I Sensaciones*, trad. Héctor Islas, IIF-UNAM, México, 2003.
- Nithianantharajha, Jess, y Anthony J. Hannan, “Enriched environments, experience-dependent plasticity and disorders of the nervous system”, *Nature*, vol. 7, pp. 697-709, 2006.
- O’Connor, Timothy y Hong Yu Wong, “Emergent properties”, *Stanford encyclopedia of philosophy*, <http://plato.stanford.edu/entries/properties-emergent/>, 2006.
- Oppenheim, Paul y Hilary Putnam, “Unity of science as a working hypothesis”, R. Boyd, P. Gasper y J. D. Trout (eds.), *The philosophy of science*, 1991.
- Penrose, Roger, *La mente nueva del emperador*, FCE, México, 2002.
- Putnam, Hilary, *Cerebro y conducta*, Cuadernos de Crítica #23, IIF-UNAM, 1983.
- , *La vida mental de algunas máquinas*, Cuadernos de Crítica #17, IIF-UNAM, 1981.
- , *La naturaleza de los estados mentales*, Cuadernos de Crítica #15, IIF-UNAM, 1981.
- , *Otras mentes*, Cuadernos de Crítica, #26, IIF-UNAM, 1984.
- Ridley, Mark, *Evolution*. Blackwell Publishing, Malden, 2003.
- Robinson, Howard, “Dualism”, *Stanford encyclopedia of philosophy*, <http://plato.stanford.edu/entries/dualism/>, 2007.
- Ruse, Michael, *La filosofía de la biología*, Alianza Editorial, Madrid, 1973.
- Sapp, Jan, *Genesis: The evolution of biology*, Oxford University Press, 2003.
- Searle, John. R., *El redescubrimiento de la mente*, Crítica, 1994.
- *Mentes cerebros y ciencia*, 4ª ed., trad. Luis Valdés, Catedra Teorema, 2001.

- , "Minds, brains, and programs", *Behavioral and brain sciences*, vol. 3 no. 3, pp. 417-457, 1980.
- Shapiro, Lawrence A., "Mind the evolution", Denis M. Walsh, (ed.) *Naturalism, evolution and mind*, Cambridge University Press, pp. 23-41, 2001
- Shoemaker, Sydney, "Absent qualia are impossible-A reply to Block", *The philosophical review*, vol. 90, no. 4, pp. 581-599, 1981.
- , "Funcionalismo y qualia", *La naturaleza de la experiencia. Vol. I Sensaciones*. Maite Ezcurdia y Olbeth Hansberg (comps.), trad. Ana Isabel Stellino, IIF-UNAM, México, 2003.
- , "Some varieties of functionalism", *Identity, cause and mind, philosophical essays*, Cambridge, 1984.
- Smart, J. J. C., "Sensations and brain processes", *Philosophical review*, no. 68, pp. 141-156, 1959.
- , "The Identity Theory of Mind", *Stanford encyclopedia of philosophy*, <http://plato.stanford.edu/entries/mind-identity/>, 2007.
- Sober, Elliot, *Philosophy of biology. Dimensions of philosophy*, University of Wisconsin-Madison, 2000.
- , *Reconstructing the past*, MIT Press, Cambridge, 1988.
- Sterelny, Kim y Paul, Griffiths, *Sex and death*. The University of Chicago Press, Chicago, 1999.
- Stoljar, Daniel, "Physicalism", *Stanford encyclopedia of philosophy*, <http://plato.stanford.edu/entries/physicalism/>, 2001.
- Turing, Alan, M., "On computable numbers, with an application to the Entscheidungsproblem", Jack Copeland (ed.), *The essential Turing: Seminal writings in computing, logic, philosophy, artificial intelligence, and artificial life: Plus the secrets of enigma*, Oxford University Press, 1936.