
LA NATURALEZA
DE LA EXPERIENCIA

VOLUMEN I:
SENSACIONES

Compilación, introducción
y revisión técnica de las traducciones

MAITE EZCURDIA
y
OLBETH HANSBERG

UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
MÉXICO 2003

Agradecemos a los proyectos CONACYT 3128P-H “Conciencia y estados mentales”, y CONACYT 27748-H “Los límites de la conciencia”, el apoyo financiero para la traducción y elaboración de esta antología. Asimismo, agradecemos a los participantes en dichos proyectos las discusiones sobre el tema de sensaciones y *qualia*, a partir de las cuales se llevó a cabo esta selección.

DR © 2003, Universidad Nacional Autónoma de México

INSTITUTO DE INVESTIGACIONES FILOSÓFICAS

Circuito Mtro. Mario de la Cueva s/n,
Ciudad Universitaria, Coyoacán, 04510, México, D.F.
Tels.: 56 22 74 37 y 56 22 75 04; fax 56 65 49 91
correo electrónico: libros@filosoficas.unam.mx
página web: <http://www.filosoficas.unam.mx>

Impreso y hecho en México
ISBN 970-32-0546-1

LO QUE NO SON LOS ESTADOS PSICOLÓGICOS*

NED BLOCK Y JERRY FODOR

I

Hasta donde se sabe, organismos diferentes a menudo se hallan en estados psicológicos de exactamente el mismo tipo en uno u otro momento, y un organismo dado se encuentra a menudo en estados psicológicos de exactamente el mismo tipo en momentos diversos. Siempre que se dé uno de estos casos, diremos de los estados psicológicos del organismo o de los organismos en cuestión que son de *tipo idéntico*.

Una de las tareas que frecuentemente intentan realizar las teorías de la filosofía de la mente que hoy en día están en boga es caracterizar las condiciones para la identidad tipo de los estados psicológicos. Por ejemplo, ciertas versiones del conductismo filosófico afirman que dos organismos se encuentran en estados psicológicos de tipo idéntico si y sólo si algunas de sus conductas o disposiciones conductuales son de tipo idéntico. En forma análoga, algunas (aunque no todas) las versiones del fisicalismo aseveran que los organismos se encuentran en estados

* Varios amigos y colegas han leído los primeros borradores. Estamos particularmente en deuda con los profesores Richard Boyd, Donald Davidson, Michael Harnish y Hilary Putnam por la atención con que leyeron el escrito y por sus útiles sugerencias.

psicológicos de tipo idéntico si y sólo si algunos de sus estados físicos son de tipo idéntico.¹

Mientras se interpreten como teorías sobre las condiciones de identidad tipo de los estados psicológicos, parece cada vez más improbable que el conductismo o el fisicalismo sean verdaderos. Ya que los argumentos a favor de esta conclusión están presentes en una amplia gama de literatura, expondremos aquí sólo una brevísima revisión.²

El argumento fundamental contra el conductismo consiste, sencillamente, en que lo que un organismo hace o está dispuesto a hacer en un momento dado es una función muy compleja de sus creencias y deseos, junto con sus entradas sensoriales presentes y sus recuerdos. Es, por lo tanto, extremadamente improbable que resulte posible casar predicados conductuales con predicados psicológicos en la forma exigida por el conductismo; a saber, que para cada tipo de estado psicológico, un organismo se halla en ese estado si y sólo si un predicado conductual específico es verdadero de él. Esto sugiere que es altamente probable que el conductismo sea falso simplemente en virtud de sus consecuencias empíricas e independientemente de su poca plausibilidad como tesis semántica. El conductismo no puede ser verdadero a menos que el correlacionismo entre la mente y la conducta sea verdadero, y el correlacionismo entre la mente y la conducta no es verdadero.

El argumento contra el fisicalismo descansa en la probabilidad empírica de que criaturas de composición y estructu-

¹ Si el fisicalismo es la doctrina de que los estados psicológicos son estados físicos, tenemos entonces dos versiones, dependiendo de si interpretamos a "estados" como refiriéndose a tipos o a casos. La última interpretación ofrece una teoría más débil al suponer que un caso del tipo *x* puede ser idéntico a un caso del tipo *y* a pesar de que *x* y *y* sean tipos distintos. Según este supuesto, el fisicalismo tipo implica claramente el fisicalismo caso, pero no a la inversa.

La distinción entre teorías de identidad caso y las teorías de identidad tipo no ha sido explotada en el caso de los análisis conductuales. Al contrario de ambas versiones del fisicalismo, el conductismo se sostiene por lo general como una tesis semántica, por lo tanto, como una teoría sobre las relaciones lógicas entre tipos. En este artículo, "fisicalismo" significará fisicalismo tipo. Al hablar de los estados especificaremos si nos estamos refiriendo a tipos o casos sólo cuando no resulte claro a partir del contexto.

² Véase Davidson 1970; Fodor 1968; Putnam 1965; 1966; 1967.

ra diferentes, que no están en un sentido interesante en estados físicos idénticos, puedan no obstante estar en estados psicológicos idénticos; por ende, de que los tipos de estados psicológicos no guarden correspondencia con los tipos de estados físicos. Este argumento ha sido desarrollado convincentemente en Putnam 1967. En esencia, recurre a las tres siguientes clases de consideraciones empíricas.

En primer lugar, la doctrina de Lashley sobre la equipotencialidad neurológica sostiene que cualquiera de una amplia gama de funciones psicológicas puede ser cumplida mediante cualquiera de una amplia gama de estructuras cerebrales. Si bien la generalidad de esta doctrina es cuestionable, sí parece claro que el sistema nervioso central es sumamente versátil y que un tipo dado de proceso psicológico está frecuentemente asociado con una variedad de estructuras neurológicas distintas. (Por ejemplo, es un hecho ampliamente conocido que un trauma a edad temprana puede conducir al establecimiento de funciones lingüísticas en el hemisferio *derecho* de personas diestras.) Pero el fisicalismo, según lo hemos entendido, requiere que los organismos estén en estados psicológicos de tipo idéntico si y sólo si están en estados físicos de tipo idéntico. Por ende, si la equipotencialidad es verdadera, el fisicalismo tiene que ser falso.

La segunda consideración depende del supuesto de que la doctrina darwiniana de la convergencia se aplica a la filogenia de la psicología, así como a la filogenia de la morfología y del comportamiento. Es bien sabido que similitudes morfológicas superficiales entre los organismos pueden no representar más que soluciones evolutivas paralelas al mismo problema ambiental: en particular, que pueden ser la expresión de tipos completamente distintos de estructuras fisiológicas. La idea análoga acerca de las similitudes conductuales entre especies ha sido ampliamente reconocida en la literatura etológica: organismos de filogenia y morfología muy diferentes pueden llegar, no obstante, a exhibir similitudes conductuales superficiales en respuesta a presiones ambientales convergentes. La idea que nos interesa aquí es que bien pueden aplicarse las mismas consideraciones a la filogenia de la psicología de los organismos. Las similitudes psicológicas entre las especies pueden reflejar a menudo

una selección ambiental convergente en vez de similitudes fisiológicas subyacentes. Por ejemplo, carecemos de una razón especial para suponer que la fisiología del dolor en el hombre deba tener mucho en común con la fisiología del dolor en especies filogenéticamente alejadas de él. Pero si existen organismos cuya psicología es homóloga a la nuestra, pero cuya fisiología es totalmente distinta, tales organismos ofrecen contraejemplos a las correlaciones psicofísicas que requiere el fisicalismo.

Por último, si admitimos la posibilidad conceptual de que los predicados psicológicos puedan aplicarse a los artefactos, parece entonces probable que el fisicalismo resulte empíricamente falso. Pues parece probable que, dada una correlación psicofísica cualquiera que vale para un organismo, es posible fabricar una máquina que sea psicológicamente semejante a ese organismo, pero que sea fisiológicamente lo suficientemente distinta de ese organismo como para que la correlación psicofísica no valga para la máquina.

Lo que estos argumentos parecen mostrar es que las condiciones que el conductismo y el fisicalismo intentan imponer sobre la identidad tipo de los estados psicológicos de los organismos son, en un sentido pertinente, insuficientemente abstractas. Parece probable que organismos que difieren en su comportamiento o en sus disposiciones conductuales puedan hallarse no obstante en estados psicológicos de tipo idéntico, al igual que pueden estarlo organismos en estados físicos distintos. (Discutiremos más adelante un punto de vista "funcionalista" sobre la identidad tipo, el cual intenta fijar los criterios de identidad en un nivel más abstracto del que reconocen el fisicalismo o el conductismo.)

Desde luego, es *posible* que resulte que se dan las correspondencias entre tipos exigidas por el conductismo o por el fisicalismo. El argumento aquí es que, incluso si los estados conductuales o físicos *están* en correspondencia uno a uno con los estados psicológicos, no tenemos hasta ahora ninguna evidencia de que esto sea así; por lo tanto, carecemos de justificación para adoptar teorías filosóficas que lo *exijan*. La paradoja del conductismo y el fisicalismo consiste en que, mientras que la mayor parte de los argumentos que rodean a estas doctrinas

son estrictamente “conceptuales”, parece cada vez más probable que las refutaciones decisivas de las mismas sean empíricas.

Suele sugerirse que se podrá responder a estos argumentos suponiendo que, aunque ni los estados conductuales ni los físicos se corresponden uno a uno con los estados psicológicos, puede empero haber una correspondencia de muchos a uno. Esto es, se supone que para cada tipo de estado psicológico existe una disyunción distinta de tipos de estados conductuales (o físicos) tal que un organismo está en ese estado psicológico si y sólo si está en uno de los disyuntos.

No obstante, esta clase de propuesta se topa con serias dificultades. En primer lugar, no es obvio que haya, de hecho, una disyunción *distinta* de estados conductuales (o físicos) que corresponda con cada estado psicológico. Por ejemplo, no hay en realidad ninguna razón para creer que la clase de tipos de conducta que, en la historia total del universo, han expresado (o expresarán) ira en tal o cual organismo sea distinta de la clase de tipos de conductas que han expresado, digamos, dolor. Al considerar esta posibilidad, se debe tener en cuenta que casi cualquier conducta puede convertirse, en las circunstancias adecuadas, en la expresión convencional de prácticamente cualquier estado psicológico, y que un organismo dado que se encuentra en un estado psicológico dado puede exhibir prácticamente cualquier disposición conductual, dependiendo de sus creencias y preferencias. La misma clase de razonamiento sirve, *mutatis mutandis*, para refutar el supuesto de que existe una disyunción distinta de tipos de estados físicos para cada tipo de estado psicológico, pues parece plausible que prácticamente cualquier tipo de estado físico pueda realizar prácticamente cualquier tipo de estado psicológico en una u otra clase de sistema físico.

Pero incluso si *hay* realmente una disyunción distinta de tipos de estados conductuales (o físicos) que corresponda con cada tipo de estado psicológico, no habría ninguna razón para creer que tal correspondencia es legaliforme; y no queda claro qué interés filosófico habría en el descubrimiento de una propiedad conductual (o física) que accidentalmente resultara ser coextensiva con un predicado psicológico. Así, como Davidson ha señalado, según el supuesto de que las correlaciones psicoconductuales no son legaliformes, aunque “encontráramos una ora-

ción abierta formulada en términos conductuales y exactamente coextensiva con un predicado mental, no podría razonablemente convencernos de que la hemos encontrado” (Davidson 1970). Y como también lo señaló Davidson, la misma observación se aplica, *mutatis mutandis*, a los predicados físicos.

Finalmente, una teoría que diga que cada predicado psicológico es coextensivo con una disyunción distinta de predicados conductuales (o físicos)³ es incompatible con lo que hemos supuesto que es una verdad evidente, a saber, que un determinado estado conductual puede expresar (o un determinado estado físico puede realizar) diferentes estados psicológicos en diferentes momentos. Supongamos, por ejemplo, que tenemos una teoría que establece que el predicado psicológico p_1 es coextensivo con el predicado conductual disyuntivo α y que el predicado psicológico p_2 es coextensivo con el predicado conductual disyuntivo β . Supongamos, además, que S_i designa un tipo de conducta que en algunas ocasiones ha expresado p_1 pero no p_2 , y que en otras ha expresado p_2 pero no p_1 . Así pues, S_i tendrá que ser un disyunto tanto de α como de β . Pero los disyuntos de α son cada uno condiciones suficientes de p_1 y los disyuntos de β son cada uno condiciones suficientes de p_2 , bajo el supuesto de que p_1 y α , y p_2 y β , son, respectivamente, coextensivos. Luego, la teoría implica que un organismo en S_i está tanto en p_1 como en p_2 , lo cual es lógicamente incompatible con la afirmación de que S_i expresa p_1 (pero no p_2) en ciertas ocasiones y p_2 (pero no p_1) en otras. Desde luego, se podría eludir esta objeción incluyendo designadores espaciotemporales en la especificación de los disyuntos mencionados en α y β . Pero hacer esto sería abandonar por completo el proyecto de expresar correlaciones psicoconductuales (o psicofísicas) mediante bicondicionales legaliformes.

³ No todos los conductistas filosóficos mantienen esta posición; el conductismo filosófico puede ser caracterizado en términos generales como la posición de que para cada predicado psicológico existe un predicado conductual con el que guarda una “relación lógica”. (Véase Fodor 1968.) Luego, la siguiente idea califica como conductista: todas las adscripciones de predicados psicológicos implican adscripciones de predicados conductuales, pero no a la inversa. Aunque esta forma de conductismo no es vulnerable al presente argumento, los anteriores la rebaten con tanta eficacia como a las formas bicondicionales del conductismo.

II

Se ha propuesto recientemente que es posible sortear esta clase de dificultades y formular una teoría adecuada acerca de las condiciones para la identidad tipo de los estados psicológicos de la siguiente manera: supongamos que cualquier sistema P , al cual pueden aplicarse predicados psicológicos, puede describirse como un autómata probabilístico. (Un autómata probabilístico es una máquina de Turing generalizada, cuya tabla de máquina contiene instrucciones asociadas con probabilidades positivas finitas que son menores que o iguales a uno. Para una breve introducción al concepto de una máquina de Turing, una tabla de máquina y otras nociones afines, véase Putnam 1967.) Una *descripción* de P , en el sentido técnico que aquí pretende dársele, es cualquier enunciado verdadero de que P posee distintos estados S_1, S_2, \dots, S_n , los cuales están relacionados entre sí y con las salidas y entradas de P por medio de las probabilidades de transición dadas en una tabla de máquina específica. Llamaremos a los estados S_1, S_2, \dots, S_n , especificados por la *descripción* de un organismo, los “estados de la tabla de máquina del organismo” correspondientes a esa *descripción*.

Es en el contexto del supuesto de que los organismos pueden ser descritos como autómatas, que esta teoría (en adelante “TIEF”, por “teoría de identidad de los estados funcionales”) busca especificar las condiciones para la identidad tipo de los estados psicológicos. En particular, la TIEF afirma que, para cualquier organismo que satisfaga predicados psicológicos, existe una *descripción* única óptima, conforme a la cual cada estado psicológico del organismo es idéntico a uno de sus estados de tabla de máquina relativos a esa descripción.

Son pertinentes varias observaciones acerca de la TIEF. En primer lugar, hay una generalización obvia de la noción de autómata probabilístico, en la cual se considera que tiene una cinta de entradas independiente sobre la que un “oráculo” puede imprimir símbolos durante un proceso computacional. La TIEF presupone una interpretación de esta generalización, según la cual unos transductores sensoriales toman el lugar del “oráculo” y las salidas se consideran como instrucciones a los transductores motrices. Una interpretación semejante debe esperarse si

una *descripción* de un organismo ha de proporcionar un modelo de las operaciones mentales del organismo.

En segundo lugar, hemos presentado la TIEF de la manera habitual, como una teoría de *identidad*:⁴ en particular, como una teoría que afirma que cada tipo de estado psicológico es idéntico a (un tipo de) estado de tabla de máquina. Sin embargo, nuestro objetivo es dejar de lado las interrogantes acerca de las condiciones de identidad de los objetos abstractos y discutir sólo cierta clase de bicondicionales que los enunciados de identidad tipo-tipo implican: esto es, enunciados de la forma “*O* presenta tal y cual tipo de estado psicológico en el tiempo *t* si y sólo si *O* está en tal y cual tipo de estado de tabla de máquina en el tiempo *t*”.

En tercer término, vale la pena insistir en que la TIEF es más que la afirmación de que todo organismo posee una descripción como máquina de Turing o como autómata probabilístico. Pues existen varios sentidos en los cuales esa afirmación es trivialmente verdadera; pero su verdad en estos sentidos no implica la TIEF. Por ejemplo, si las entradas y salidas de un organismo son recursivamente enumerables (como sucede con cualquier organismo mortal después de su fallecimiento), se sigue entonces que existe una máquina de Turing capaz de simular al organismo (es decir, una máquina de Turing que tiene las mismas relaciones entre entradas y salidas). Pero no se sigue que el organismo posee una *descripción* única óptima del tipo caracterizado más arriba. En segundo lugar, como lo ha señalado Putnam (en conversación), *todo* es describable como una realización de lo que podría llamarse máquina de Turing “nula”, esto es, una máquina que sólo posee un estado y permanece en él. (El argumento consiste, toscamente, en que el que un sistema *P* realice o no una máquina de Turing depende, entre otras cosas, de lo que cuenta como un cambio de estado en *P*. Si se hace valer *nada* como un cambio de estado en *P*, entonces *P* constituye una realización de la máquina de Turing nula.) Pero nuevamente, la TIEF no sería verdadera si la única *descripción* verdadera de un organismo fuera una máquina de Turing nula, ya que la TIEF requiere que los estados de tabla de máquina de un organismo

⁴ Cfr. Putnam 1967 y 1970.

correspondan uno a uno con sus estados psicológicos bajo su descripción óptima.

Hay así dos sentidos importantes en los cuales la TIEF involucra más que la afirmación de que los organismos que satisfacen predicados psicológicos poseen descripciones. En primer lugar, la TIEF afirma que tales sistemas poseen descripciones únicas óptimas. En segundo lugar, la TIEF afirma que los tipos de estados de tabla de máquina especificados por la descripción única óptima de un sistema corresponden a los tipos de estados psicológicos en que puede estar el sistema. Es esta segunda afirmación de la TIEF la que constituirá nuestro principal interés.

La TIEF, a diferencia del conductismo o del fisicalismo, no es una teoría ontológica: esto es, es neutral con respecto a lo que *son* los estados psicológicos caso, por cuanto, hasta donde le interesa a la TIEF, entre los sistemas a los cuales se aplican los predicados psicológicos (y que en consecuencia poseen *descripciones*) pueden incluirse personas, objetos materiales, almas y otros. Esta última idea indica cómo la TIEF es capaz de soslayar algunos de los tipos de dificultades que planteamos con respecto al fisicalismo y el conductismo. Así como la TIEF se abstrae de las consideraciones acerca del estatus ontológico de los sistemas que poseen *descripciones*, también se abstrae de las diferencias físicas entre sistemas que poseen *descripciones* comunes. Como lo ha observado Putnam, “la *misma* máquina de Turing (desde el punto de vista de la tabla de máquina) puede realizarse físicamente en una variedad potencialmente infinita de formas” (1966, p. 271), y la TIEF nos permite enunciar condiciones para la identidad tipo de estados psicológicos que son neutrales respecto de esas realizaciones diferentes.

Asimismo, la TIEF nos permite enunciar dichas condiciones en términos que se abstraen de la variedad de consecuencias conductuales que pueda tener un estado psicológico. De este modo, supera un tipo de objeción que, como supusimos más arriba, era fatal para el conductismo.

Observamos que el conductista está comprometido con la idea de que dos organismos se encuentran en el mismo estado psicológico siempre que sus conductas y/o disposiciones conductuales sean idénticas; y que esta teoría es poco plausible en

la medida en que las conductas y las disposiciones conductuales de un organismo son efectos de las *interacciones* entre sus estados psicológicos. Pero la TIEF nos permite distinguir entre los estados psicológicos no sólo en términos de sus consecuencias conductuales, sino también en términos del carácter de sus interconexiones. Esto se debe a que el criterio de identidad para los estados de tabla de máquina reconoce *las relaciones entre ellos*, además de sus relaciones con las entradas y salidas. Así, la TIEF puede hacer frente al carácter indirecto distintivo de la relación entre estados psicológicos y comportamiento. De hecho, la TIEF nos permite ver cómo estados psicológicos que *no* tienen una expresión conductual pueden ser, no obstante, distintos.

Finalmente, cabe notar que nada impide tomar por lo menos algunas de las transiciones especificadas en una tabla de máquina como si correspondiera a las relaciones causales del sistema que la tabla *describe*. En particular, puesto que la TIEF es compatible con el fisicalismo caso, no hay motivos para que no reconozca que los estados psicológicos caso pueden entrar en relaciones causales. Así, cualquier ventaja que redunde en beneficio de los análisis causales de los estados psicológicos, o de las relaciones entre estados psicológicos y conducta, redundará igualmente en beneficio de la TIEF.⁵

III

En esta sección vamos a revisar una serie de argumentos que ofrecen cierto grado de dificultad a la afirmación de que la TIEF proporciona una explicación adecuada de las condiciones para la identidad tipo de los estados psicológicos. Nuestro punto de vista es que, tomados en conjunto, estos argumentos son suficientemente decisivos contra la teoría de identidad tipo de los estados psicológicos propuesta por la TIEF. En la sección final indicaremos algunas razones acerca de por qué el intento por ofrecer condiciones sustanciales para la identidad tipo de los estados psicológicos fracasa con tanta frecuencia.

(1) Cualquier explicación de las condiciones para la identidad tipo de los estados psicológicos que pretenda adherirse

⁵ Cfr. Davidson 1963.

estrechamente a nuestra noción cotidiana de qué tipos de estados psicológicos hay seguramente tendrá que establecer una distinción entre estados disposicionales (creencias, deseos, inclinaciones, etc.) y estados incidentes* (sensaciones, pensamientos, emociones, etc.). Sin embargo, hasta donde se alcanza a ver, la TIEF carece de una vía plausible para captar esta distinción a menos que abandone su principio fundamental de que los estados psicológicos se corresponden uno a uno con los estados de la tabla de máquina. Supongamos, por ejemplo, que la TIEF intenta reconstruir la distinción entre incidentes y disposiciones haciendo referencia a la distinción entre el estado de la tabla de máquina *en* el que está un organismo y todos los otros estados especificados por su tabla de máquina. Así, se podría pulir la TIEF de suerte que dijera: para los estados incidentes, dos organismos se encuentran en estados psicológicos de tipo idéntico si y sólo si están en el mismo estado de tabla de máquina; y para cada estado disposicional, existe un estado de tabla de máquina tal que un organismo está en el primero si y sólo si su tabla de máquina contiene el segundo.

Lo siguiente indica una forma de poner de manifiesto la poca plausibilidad de esta propuesta. Todo estado de tabla de máquina de un organismo es un estado en el que ese organismo puede estar en uno u otro momento. Por lo tanto, si la distinción entre el estado de tabla de máquina en el que está un organismo y todos los otros estados de su tabla de máquina es igual a la distinción entre el estado incidente de un organismo y sus estados disposicionales, se sigue que todo estado disposicional del organismo es un estado incidente posible de ese organismo.

Esta consecuencia de la TIEF es correcta para numerosas clases de disposiciones psicológicas. Por ejemplo, en correspondencia con el predicado disposicional “habla francés”, tenemos el predicado incidente “está hablando francés”; en correspondencia con el predicado disposicional “es voraz”, tenemos el predicado incidente “está siendo voraz”; en correspondencia con el predicado disposicional “puede oír sonidos mayores de 3 000 hercios”, tenemos el predicado incidente “está oyendo un

* Hemos traducido “*ocurrent*” como “incidente” tanto en su forma adjetival como sustantiva. [N. de las comps.]

sonido mayor a 3 000 hercios". Y, en general, para muchos predicados disposicionales, tenemos las formas correspondientes en tiempo presente continuo que denotan incidencias.

Sin embargo, para muchos otros predicados disposicionales psicológicos, este paralelismo no se sostiene. Por ejemplo, carecemos de un predicado "está creyendo que *P*" que corresponda a "cree que *P*"; carecemos de "está deseando un caramelo" que corresponda con "desea un caramelo"; carecemos de "está prefiriendo *X* sobre *Y*" que corresponda con "prefiere *X* a *Y*", y así sucesivamente. En suma, muchos estados psicológicos disposicionales *no* son estados psicológicos incidentes posibles, y para éstos la TIEF no ofrece un modelo claro.

Es importante entender lo que este argumento *no* demuestra. De acuerdo con este argumento, ciertos estados disposicionales no pueden corresponder a estados de tabla de máquina porque todos los estados de tabla de máquina son estados incidentes posibles, pero algunos estados psicológicos disposicionales no lo son. Para esas disposiciones, no puede haber estados en la tabla de máquina que permitan afirmar que el organismo tiene el primero si y sólo si aparece el segundo en su *descripción*. Sin embargo, es perfectamente posible que las condiciones necesarias y suficientes para tener tales disposiciones deban darse por referencia a alguna propiedad *abstracta* de la organización de las tablas de máquina. Tomando un ejemplo improbable, dada una forma normal de las descripciones, podría resultar que un organismo creyera que el Sol se encuentra a 149.5 millones de kilómetros de la Tierra si y sólo si las primeras *n* columnas de su tabla de máquina poseen una propiedad abstracta tal que contiene únicamente números enteros nones. Puesto que decir de una máquina que las primeras *n* columnas... , etc., no es adscribirle un estado de tabla de máquina, los estados psicológicos que se analizan como correspondientes a esta clase de propiedad no se describirían por ello como estados incidentes posibles.

Sin embargo, seguir este camino implicaría abandonar una afirmación fundamental de la TIEF. Pues, aunque esta manera de abordar el asunto resulta compatible con la idea de que dos organismos tienen la misma psicología si y sólo si tienen la misma tabla de máquina, no es compatible con la sugerencia

de que dos organismos se encuentran en el mismo estado psicológico (disposicional) si y sólo si tienen en común un estado especificado en sus tablas de máquina. Por ende, es incompatible con la idea de que los estados psicológicos corresponden uno a uno con los estados de tabla de máquina. Además, como no tenemos manera de saber qué clases de propiedades abstractas de las tablas de máquina podrían llegar a corresponder con los estados psicológicos, la dificultad presente reduce en gran medida la posibilidad de usar la TIEF para delinear condiciones sustanciales para la identidad tipo de los estados psicológicos. Decir que los estados psicológicos corresponden a una u otra propiedad de las tablas de máquina es decir algo mucho más débil que decir que los estados psicológicos corresponden a estados de la tabla de máquina. Este es un tipo de argumento al que volveremos más adelante en la discusión.

Hay, desde luego, por lo menos otro camino para salir de la presente dificultad de la TIEF. Podría sugerirse que debemos abandonar la noción cotidiana de que hay ciertos estados disposicionales que no son estados incidentes posibles (por ejemplo, reconocer un estado incidente, aunque tal vez no consciente, de estar creyendo que *P*). Está claro que la posibilidad de que un día podríamos tener bases teóricas para reconocer la existencia de tales estados no puede excluirse *a priori*. Pero carecemos de tales bases *en este momento*, y nos parece que sí hay un principio metodológico del conservacionismo que establece que debemos evitar los modelos que exigen cambios empíricos o conceptuales que no estén motivados de manera independiente.

(2) Sugerimos que la TIEF nos permite explicar el hecho de que la conducta es típicamente el producto de las interacciones entre estados psicológicos, y que la existencia de tales interacciones constituye una fuente constante de dificultades para las teorías conductistas, en cuanto éstas buscan asignar conductas características *individualmente* a estados psicológicos. Desde el punto de vista empírico, sin embargo, es sumamente probable que haya *dos* clases de interacciones conductualmente eficaces entre los estados psicológicos, y la TIEF sólo ofrece un modelo natural para una de ellas.

Por un lado, el comportamiento puede ser producto de una *serie* de estados psicológicos, y la explicación de la TIEF nos mues-

tra de qué manera podría resultar esto cierto y cómo algunos de los estados que ocurren en esa serie pueden ellos mismos carecer de expresiones conductuales. Pero, por otro lado, la conducta puede ser el resultado de interacciones entre estados mentales *simultáneos*. Por ejemplo, a primera vista, lo que un organismo hace en t puede ser una función de lo que siente en t y de lo que piensa en t . Pero la TIEF no proporciona ninguna herramienta conceptual que nos permita representar este estado de cosas. En efecto, la TIEF es capaz de ofrecer la representación de interacciones secuenciales entre estados psicológicos, pero no de interacciones simultáneas. De hecho, la TIEF ni siquiera da cuenta del hecho de que un organismo puede estar en más de un estado psicológico incidente al mismo tiempo, ya que un autómata probabilístico sólo puede estar en un estado de tabla de máquina a la vez. La consecuencia de este argumento parece ser que si ha de utilizarse a los autómatas probabilísticos como modelos de un organismo, el modelo adecuado es un conjunto de autómatas intercomunicados que operen en forma paralela.

También es importante tener en claro lo que este argumento no demuestra con respecto a la TIEF. Hemos interpretado la TIEF como si afirmara que los estados psicológicos de un organismo guardan una correspondencia uno a uno con los estados de tabla de máquina postulados en su *descripción* óptima. El presente argumento indica que, para aceptar esta afirmación, la *descripción* óptima de un organismo no debe representarlo como un autómata probabilístico individual. Si los organismos, pero no los autómatas probabilísticos individuales, pueden estar en más de un estado a la vez, entonces o bien un organismo no es un autómata probabilístico individual, o bien los estados psicológicos de un organismo no corresponden a los estados de tabla de máquina de los autómatas probabilísticos individuales. (Debe señalarse que existe un algoritmo para construir un autómata individual que sea equivalente a cualquier conjunto dado de autómatas paralelos. Sin embargo, no puede darse el caso de que *tanto* un conjunto de autómatas paralelos *como* el autómata individual equivalente ofrezcan a la vez *descripciones* óptimas de un organismo.)

Estas observaciones poseen cierta importancia, ya que el tipo de teoría psicológica que obtenemos a partir del supuesto de

que los organismos son procesadores paralelos resulta totalmente distinta del tipo que obtenemos a partir del supuesto de que son procesadores en serie. De hecho, aunque las características generales de los procesadores en serie se comprenden relativamente bien, se sabe muy poco de las correspondientes características de los sistemas paralelos.

Por otra parte, este argumento deja incólume la afirmación principal de la TIEF: incluso si los organismos son, en cierto sentido, conjuntos de autómatas probabilísticos, puede ocurrir que cada uno de los estados psicológicos de un organismo corresponda al estado de tabla de máquina de uno u otro miembro del conjunto. En los argumentos que siguen, asumiremos el modelo en serie por razones de simplicidad e intentaremos mostrar que, incluso bajo ese supuesto, los estados psicológicos no corresponden a los estados de la tabla de máquina.

(3) La TIEF sostiene que dos organismos están en estados psicológicos del mismo tipo si y sólo si están en el mismo estado de tabla de máquina. Pero los estados de tabla de máquina son idénticos si y sólo si están relacionados de modo idéntico con otros estados de tabla de máquina y con los estados de los mecanismos de entrada y salida. En este sentido, el criterio para la identidad de los estados de tabla de máquina es la “equivalencia funcional”. Por lo tanto, la TIEF afirma que la identidad tipo de los estados psicológicos es igualmente una cuestión de cierta clase de equivalencia funcional; los estados psicológicos son de tipo idéntico si y sólo si comparten las propiedades que deben especificarse para individuar un estado de tabla de máquina.

Sin embargo, podría argumentarse de manera plausible que este modo de identificar en tipos a los estados psicológicos no logra dar cabida a un rasgo de por lo menos algunos de esos estados, el cual es crucial para determinar su tipo: a saber, su carácter “cualitativo”. No parece, por ejemplo, totalmente irrazonable sugerir que no habría ningún caso del tipo “estado de dolor” a menos que se sintiera como dolor, y que esto sería verdad aun cuando estuviera relacionado con todos los otros estados psicológicos del organismo de cualesquiera que sean las formas en que los dolores están relacionados. Nos parece que los contraargumentos verificacionistas estándares que se oponen a la idea de que la hipótesis del “espectro invertido” es

conceptualmente incoherente son poco convincentes. Si esto es correcto, daría entonces la impresión de que la posibilidad de la inversión de los *qualia* representa un argumento *prima facie* serio contra las explicaciones funcionalistas de los criterios para la identidad tipo de los estados psicológicos.

Debe observarse, sin embargo, que el argumento de la inversión de los *qualia* constituye *sólo* una objeción *prima facie* contra la TIEF. En particular, el proponente de las explicaciones funcionalistas dispone de cualquiera de dos medios para hacer frente a esta objeción. Por un lado, podría argumentar que, aunque los *qualia* invertidos, *si se dieran*, ofrecerían contraejemplos de su teoría, es imposible, nomológicamente, que estados psicológicos funcionalmente idénticos sean cualitativamente distintos: en particular, que cualquier cosa que alterase las características cualitativas de un estado psicológico alteraría sus características funcionales. Esta línea de razonamiento podría parecer un flagrante apriorismo, pero, en ausencia de los datos empíricos pertinentes, bien puede adoptarse la actitud de esperar y ver.

Hay, además, otra defensa posible para el proponente de la TIEF. Puede decir que, dados dos estados psicológicos funcionalmente idénticos, podríamos (o quizá “deberíamos”) *tomarlos* como idénticos en tipo, independientemente de sus propiedades cualitativas: es decir, que las diferencias entre las propiedades cualitativas de los estados psicológicos que no determinen diferencias funcionales correspondientes carecen *ipso facto* de pertinencia para los objetivos de la construcción de teorías en psicología y, por lo tanto, deben ignorarse al establecer identificaciones tipo.

Para apreciar la plausibilidad de esta sugerencia, imaginemos que se descubriera que cada persona tiene, de hecho, *qualia* ligeramente distintos (o, mejor aún, *qualia* extremadamente distintos) cuando está en cualquier estado de tabla de máquina que se supone que es idéntico al dolor. Parece bastante claro que, en este caso, podría ser razonable afirmar que el carácter de los *qualia* de un organismo carece de pertinencia con respecto a si tiene un dolor o (en términos equivalentes) que los dolores se sienten de manera muy distinta en organismos diferentes.

Esta forma de argumento puede conducir, no obstante, a consecuencias vergonzosas. Hasta donde sabemos ahora, puede ser nomológicamente posible que dos estados psicológicos sean funcionalmente idénticos (es decir, que estén relacionados idénticamente con las entradas, las salidas y los estados sucesores), aun cuando sólo uno de los estados posea un contenido cualitativo. En este caso, la TIEF nos obligaría a decir que un organismo puede tener un dolor aunque no esté sintiendo *absolutamente nada*, y esta consecuencia parece totalmente inaceptable.

Cabe observar que estos casos de “*qualia* invertidos (o ausentes)” le plantean, en cierto sentido, un problema más profundo a la TIEF del que le plantean cualquiera de los otros argumentos que discutiremos. Nuestros otros argumentos apuntan, en lo principal, a demostrar que los estados psicológicos no pueden definirse funcionalmente de cierta manera, a saber, estableciendo una correspondencia entre ellos y los estados de tabla de máquina. Pero aunque sean incompatibles con la TIEF, son compatibles con el funcionalismo en el sentido amplio de esa doctrina, la cual sostiene que las condiciones para la identidad tipo de los estados psicológicos refieren sólo a sus relaciones con las entradas, las salidas, y a las relaciones entre sí. Sin embargo, podría tomarse esta consideración como si demostrara que los estados psicológicos no pueden definirse *en lo absoluto* en términos funcionalistas, y que no pueden corresponder a *ninguna* propiedad definible en términos de autómatas abstractos. Haremos caso omiso de esta posibilidad en lo que sigue, ya que si los estados psicológicos no son estados funcionales en lo absoluto, sencillamente no surge la pregunta de si son o no estados de tabla de máquina.

(4) Observamos que hay ciertas objeciones al conductismo y al fisicalismo que indican que cada uno de éstos propone restricciones a las condiciones para la identidad tipo de los estados psicológicos que, de una u otra forma, son insuficientemente abstractas. Argumentaremos ahora que la misma clase de objeción se aplica a la TIEF.

Una tabla de máquina especifica un estado en términos de un conjunto de instrucciones que controlan el comportamiento de la máquina siempre que se encuentre en ese estado. Por

definición, en el caso de un autómata determinista, sus instrucciones especifican, para cada estado de la máquina, una salida asociada y un estado de máquina sucesor. El autómata probabilístico difiere del anterior sólo en que cualquier estado puede especificar una *gama* de salidas o de estados sucesores, con una distribución de probabilidades asociada. En suma, dos estados de tabla de máquina de un autómata determinista son distintos si difieren en sus salidas asociadas o en su estado sucesor asociado. De manera análoga, dos estados de tabla de máquina de un autómata probabilístico difieren, si difieren en su gama de salidas o en su gama de estados sucesores o en las distribuciones de probabilidades asociadas con alguna de estas gamas.

Si, no obstante, trasladamos esta convención para distinguir estados de tabla de máquina a la identificación tipo de los estados psicológicos, obtenemos condiciones de identidad que son, por así decirlo, demasiado finas. De este modo, por ejemplo, si tú y yo *sólo* diferimos en que tu respuesta más probable al dolor causado por pegarse en el dedo del pie es exclamar “maldición”, mientras que la mía es decir “maldita sea”, se sigue que el dolor que sientes al pegarte en el dedo del pie es de tipo distinto del que yo siento cuando me pego en el dedo del pie.

Este argumento se repite de manera vergonzosa. Para ver esto, consideremos el caso especial del autómata determinista: x y y son estados de tabla de máquina de distinto tipo de tal autómata si los estados sucesores inmediatos de x y y son de tipo distinto. Pero los estados sucesores inmediatos de x y y son de tipo distinto si *sus* estados sucesores inmediatos son de tipo distinto. De este modo, x y y son de tipo distinto si los sucesores inmediatos de sus sucesores inmediatos son de tipo distinto, y así sucesivamente. De hecho, bajo el supuesto de que existe una vía computacional que lleva de cada estado a todos los demás, dos autómatas cualesquiera que tengan menos que todos sus estados en común no tendrán ningún estado en común. Este argumento se generaliza a los autómatas probabilísticos de manera obvia.

Nuevamente, es importante observar lo que *no* demuestra este argumento. En particular, no demuestra que los estados psicológicos no puedan identificarse a nivel de tipos haciendo referencia a cierta clase de propiedades *abstractas* de los estados

de tabla de máquina. Pero, como señalamos en nuestra discusión sobre el argumento 1, decir que los estados psicológicos corresponden a tal o cual propiedad definible en los estados de tabla de máquina es decir mucho menos acerca de las condiciones para la identidad tipo de los estados psicológicos que lo que pretende decir la TIEF. Y el argumento presente sí parece mostrar que las condiciones utilizadas para la identificación tipo de los estados de tabla de máquina *per se* no pueden emplearse para la identificación tipo de los estados psicológicos. Es de suponerse que este tipo de razonamiento es el que tenía en mente Putnam, por ejemplo, cuando advierte que “la dificultad consistirá, desde luego, en pasar de los modelos de los organismos *específicos* a una forma *normal* de la descripción psicológica de los *organismos*” (Putnam 1967, p. 43). En suma, a primera vista puede parecer que la explotación de los criterios empleados para la identificación tipo de los estados de tabla de máquina ofrece a la TIEF conceptos que se encuentran precisamente en el nivel de abstracción requerido para la identificación tipo de los estados psicológicos. Pero, de hecho, esto no parece ser cierto.

(5) Nos parece que el siguiente argumento muestra que no es posible establecer una correspondencia uno a uno entre los estados psicológicos de los organismos y los estados de tabla de máquina de los organismos.

El conjunto de estados que constituyen la tabla de máquina de un autómata probabilístico es, por definición, una lista. Pero el conjunto de estados mentales de, por lo menos, algunos organismos (a saber, de personas) es, como una cuestión de hechos empíricos, productivo. En particular, haciendo abstracción de las limitaciones impuestas por la memoria y la mortalidad, que no son teóricamente pertinentes, hay una variedad infinita de estados psicológicos de cualquier persona dada que son de distinto tipo y que son nomológicamente posibles. La demostración más sencilla de que esto es cierto es que, bajo el supuesto de que hay una variedad infinita de oraciones declarativas que no son equivalentes, se pueden generar descripciones definidas de esos estados sustituyendo la *S* por oraciones en el esquema *A*:

A: “la creencia (pensamiento, deseo, esperanza, o lo que fuere) de que S ”

En suma, mientras que el conjunto de estados de tabla de una máquina de Turing puede, por definición, especificarse exhaustivamente mediante una lista, el conjunto de estados mentales de una persona puede, en el mejor de los casos, especificarse mediante una axiomatización finita.

A este argumento se podría objetar que no más de un subconjunto finito de las descripciones definidas generables a partir de la sustitución realizada en A designa de hecho las creencias (deseos, esperanzas o lo que fuere) nomológicamente posibles, y que esto es cierto *no* debido a las limitaciones, carentes de interés teórico, impuestas por la memoria y la mortalidad, sino más bien debido a la existencia de leyes psicológicas que limitan el conjunto de proposiciones creíbles (deseables, esperables, o lo que fuere) a un conjunto finito. Tomando un ejemplo improbable, se podría sostener que si se eliminan todas las proposiciones que quizá no son creíbles, como “ $2 + 2 = 17$ ”, “ $2 + 2 = 147$ ” y demás, el residuo es un conjunto finito.

Sin embargo, no existe razón alguna para suponer que esto sea verdad, y hay razones muy convincentes para creer que no lo es. Por ejemplo, el conjunto infinito de descripciones cuyos miembros son “la creencia de que $1 + 1 = 2$ ”, “la creencia de que $2 + 2 = 4$ ”, “la creencia de que $3 + 3 = 6$ ”, y demás, parecería designar un conjunto de creencias posibles de un organismo que es idealmente libre de limitaciones de memoria; para expresarlo de otro modo, el hecho de que haya enunciados aritméticos que a una persona le son nomológicamente imposibles de creer es consecuencia del carácter de la memoria de la gente, no del carácter de su representación mental de la aritmética.

Debe subrayarse nuevamente que esto pretende ser una afirmación empírica, aunque en extremo plausible. Es posible imaginar a una criatura idealmente libre de limitaciones de memoria cuya representación mental de la aritmética específica, no obstante, sólo un conjunto finito de creencias aritméticas posibles. El argumento aquí es que resulta sumamente improbable que nosotros seamos tales criaturas.

Una vez más es importante observar lo que este argumento *no* demuestra. Hagamos una distinción entre los *estados de tabla de máquina* de un autómeta y los *estados computacionales* de un autómeta. Por lo primero, entenderemos lo que hemos venido entendiendo desde el principio: estados especificados por columnas en la tabla de la máquina. Por lo segundo, entenderemos cualquier estado de la máquina que pueda caracterizarse en términos de sus entradas, salidas y/o estados de tabla de máquina. Según esta acepción, los predicados “acaba de realizar un cómputo que entraña trescientos setenta y dos estados de tabla de máquina”, o “acaba de demostrar el último teorema de Fermat”, o “acaba de teclear el símbolo número ‘iésimo’ en su vocabulario de salida” designan todos estados computacionales posibles de las máquinas.

Ahora bien, lo que este argumento parece mostrar es que no es posible establecer una correspondencia entre los estados psicológicos de un organismo y los estados de tabla de máquina de un autómeta. Lo que, desde luego, *no* demuestra es que no es posible establecer una correspondencia entre los estados psicológicos de un organismo y los *estados computacionales* de un autómeta. De hecho, una condición suficiente para la existencia de esta última correspondencia es que los estados psicológicos de un organismo deban ser enumerables.⁶

⁶ La afirmación de que los organismos son autómetas probabilísticos puede ser verdadera de manera interesante aunque la TIEF sea falsa; esto es, aun si los estados psicológicos no corresponden a los estados de la tabla de máquina. Por ejemplo, podría suceder que cierto subconjunto de estados psicológicos de un organismo correspondiera a un conjunto de estados de la tabla de máquina mediante los cuales se determinara el resto de su psicología. O podría resultar que lo que corresponde a cada estado de la tabla de máquina sea una conjunción de estados psicológicos... , etc. De hecho, aunque la afirmación de que cualquier organismo puede ser modelado por un autómeta probabilístico no sea interesante, la afirmación de que para cada organismo hay un autómeta probabilístico que constituye su modelo único óptimo es interesante. Y esta última afirmación no implica la TIEF ni es, en modo alguno, obviamente verdadera.

En suma, hay muchas maneras en que podría resultar que los organismos fueran autómetas en un sentido más interesante que en el sentido en que cada cosa es un autómeta conforme a cierta descripción. Nuestro argumento aquí consiste en que tales eventualidades, aunque serían importantes, no ofrecerían

(6) Hemos argumentado que, como el conjunto de estados de tabla de máquina de un autómata no es productivo, no se puede establecer una correspondencia entre ellos y los estados psicológicos de un organismo. Argumentaremos a continuación que, incluso si pudiera llevarse a cabo tal correspondencia, ésta necesariamente fracasaría en representar propiedades esenciales de los estados psicológicos. Parece bastante claro que existen similitudes estructurales por lo menos entre algunos estados psicológicos, y que una teoría exitosa de tales estados debe representar y explotar esas similitudes. Por ejemplo, está claro que hay una relación teóricamente importante entre el estado psicológico de alguien que cree que P y el estado psicológico de alguien que cree que $P \& Q$. El argumento aquí dice simplemente que representar los estados psicológicos como una lista (por ejemplo, como una lista de estados de tabla de máquina) no representa esta clase de relación estructural. Es preciso decir que creer que P es de alguna manera⁷ un componente de creer que $P \& Q$; pero el modelo de los estados de tabla de máquina carece de recursos conceptuales para decir esto. En particular, la noción de “es un componente de” no está definida para los estados de tabla de máquina.

Podría replicarse que esta clase de argumento no es estrictamente pertinente en las afirmaciones de la TIEF, pues sin duda es posible, en principio, que haya una correspondencia uno a uno entre los estados de tabla de máquina y los estados psicológicos, aunque el vocabulario adecuado para la individuación de los primeros no capte las relaciones estructurales entre los segundos.

A esta réplica, no obstante, se le escapa el meollo del problema. A fin de entender esto, consideremos el asunto mediante oraciones. La razón de que haya paralelismos estructurales entre las oraciones es que éstas se construyen a partir de un

las condiciones generales para la identificación tipo de estados psicológicos de la manera en que la TIEF pretende hacerlo.

⁷ Ciertamente “de alguna manera”. Es obvio que creer que p no es un componente de creer que $p \vee q$ del mismo modo en que creer que p es un componente de creer que $p \& q$. Igualmente obvio es que existe cierta relación entre creer que p y creer que $p \vee q$, y una teoría de las creencias tendrá que determinar cuál es esa relación.

conjunto fijo de elementos de un vocabulario por medio de la aplicación iterada de un conjunto fijo de reglas, y el vocabulario teórico requerido para analizar los modos en que las oraciones son estructuralmente semejantes es precisamente el vocabulario requerido para especificar el dominio de esas reglas. En particular, las oraciones que son estructuralmente similares comparten el léxico o las formas de derivación, o ambos. Así, las similitudes estructurales entre oraciones se explican de la misma manera en que se explica su productividad, a saber, describiéndolas como un conjunto generado en vez de a una lista.

Nuestro argumento es que las mismas consideraciones se aplican al conjunto de estados psicológicos de un organismo. Casi seguramente, también éstos son, o por lo menos incluyen, un conjunto generado, y sus similitudes estructurales corresponden, al menos en parte, a similitudes en su derivación; es decir, con los estados psicológicos, al igual que con las oraciones, el hecho de que sean productivos y el hecho de que exhiban una estructura interna son dos aspectos del mismo fenómeno. Si esto es cierto, entonces una teoría que no capte las relaciones estructurales que se dan dentro de los estados psicológicos y entre ellos no tiene prácticamente ninguna posibilidad de llegar a una *descripción* adecuada para los propósitos de la psicología teórica.

Así, este argumento, al igual que el argumento 5, conduce a la conclusión de que, si queremos pensar en la psicología de los organismos como representada por autómatas, entonces los estados psicológicos de los organismos parecen ser análogos a los estados computacionales de un autómata, y no a los estados de tabla de máquina.

IV

Hemos estado considerando teorías de la filosofía de la mente que pueden interpretarse como intentos por establecer condiciones sustanciales para la identidad tipo de los estados psicológicos. Hemos argumentado que ninguna de las principales teorías actuales parece tener éxito en esta tarea. Podría resultar aconsejable, por lo tanto, reconsiderar toda la empresa.

Supongamos que alguien quisiera saber cuáles son los criterios para identidad tipo de las entidades físicas fundamentales. Quizá la mejor forma de responder a esto es decir que dos entidades son de tipo idéntico si no difieren con respecto a ninguna magnitud física fundamental. Así, hasta donde sabemos, las condiciones para la identificación tipo de las partículas físicas elementales no refieren a su distancia del Polo Norte, aunque sí refieren a su carga. Pero observemos que esto es simplemente una consecuencia del hecho de que no existe ninguna ley física fundamental que opere sobre las entidades en función de su distancia del Polo Norte, y que *sí existen* leyes físicas fundamentales que operan sobre entidades en función de su carga.

Se podría decir que la condición básica para la identidad tipo en la ciencia es que posibilite la articulación del dominio de las leyes. Este principio se sostiene en cada nivel de descripción científica. Así, lo que resulta *pertinente* para la pregunta de si dos entidades de un nivel son de tipo distinto es el carácter de las leyes que operan sobre entidades en ese nivel. Pero si esto es lo que en general ocurre, parecería entonces que las condiciones sustanciales para la identidad tipo de los estados psicológicos se fijarán con referencia a las leyes psicológicas (y quizá neurológicas) que operan sobre esos estados, y no de ninguna otra manera.

A la luz de estas observaciones, podemos obtener una visión más clara de cuál ha sido el error con las distintas clases de teorías filosóficas que hemos venido rechazando. Por ejemplo, cabría pensar que el conductismo entraña un intento por identificar a nivel tipo estados psicológicos sólo por referencia a las leyes, cualesquiera que sean, que determinan sus efectos *conductuales*. Pero, incluso a primera vista, esto parecería un error, ya que debe haber leyes que gobiernen la interacción de los estados psicológicos, y no hay ninguna razón para creer (y muchas razones para no creer) que los estados psicológicos que se comportan uniformemente según las leyes de la primera clase se suelen comportar uniformemente según las leyes de la segunda clase.

De modo análogo, el error del fisicalismo ha sido suponer que los estados psicológicos que son distintos en su conduc-

ta según las leyes neurológicas son *ipso facto* distintos en su conducta según las leyes psicológicas. Sin embargo, todas las probabilidades indican que estados neurológicos distintos pueden ser funcionalmente idénticos. Esto es, satisfacer los criterios para la distinción tipo de estados neurológicos probablemente no garantice la satisfacción de los criterios para la distinción tipo de estados psicológicos, ni viceversa.

En suma, el problema fundamental con el conductismo y el fisicalismo radica en que la identidad tipo se determina en relación con, en el mejor de los casos, un subconjunto de leyes, que se debe suponer, se aplican a los estados psicológicos. La única justificación de esta restricción parece encontrarse en las inclinaciones reduccionistas de estas posturas. Una vez que se ha cuestionado el reduccionismo, podemos ver que las exigencias nomológicas para la identificación tipo de los estados psicológicos son probablemente sumamente complejas y variadas. Incluso lo poco que sabemos sobre las leyes psicológicas hacen parecer poco plausible que éstas reconozcan límites de tipos entre los estados psicológicos en los sitios donde los fisicalistas o los conductistas quieren demarcarlos.

La falla básica de la TIEF es, en ciertos aspectos, análoga a la del conductismo y el fisicalismo. Desde luego, la TIEF no es reduccionista, ni siquiera en espíritu, y en tanto se trata de una especie de funcionalismo, nos invita a identificar los estados psicológicos en tipos haciendo referencia a sus conexiones nomológicas con las entradas sensoriales, las salidas conductuales y las relaciones entre sí. Pero la TIEF pretende fijar una restricción adicional a la identidad tipo: a saber, que se puede establecer una correspondencia (de hecho, una identificación) entre los estados psicológicos de un organismo y los estados de tabla de máquina especificados por la *descripción* óptima del organismo. Hemos argumentado que ésta es, de hecho, una restricción sustancial, y que no puede ser satisfecha.

Lo que aparentemente queda de la TIEF es lo siguiente. Puede ser a la vez cierto e importante que los organismos sean autómatas probabilísticos. Pero incluso si es cierto e impor-

tante, el hecho de que los organismos sean autómatas probabilísticos parece tener poco o nada que ver con las condiciones para la identidad tipo de sus estados psicológicos.

[Traducción de Ana Isabel Stellino]

BIBLIOGRAFÍA

- Butler, R.J. (comp.), 1965, *Analytical Philosophy. Second Series*, Blackwell, Oxford.
- Capitan, W.H. y D.D. Merrill, 1967, *Art, Mind and Religion*, Pittsburgh University Press, Pittsburgh.
- Davidson, D., 1970, "Mental Events", en Foster y Swanson 1970, pp. 79-101. [Versión en castellano: "Sucesos mentales" en D. Davidson, *Ensayos sobre acciones y sucesos*, trad. O. Hansberg, J.A. Robles y M.M. Valdés, Crítica/Instituto de Investigaciones Filosóficas-UNAM, Barcelona, 1995, pp. 263-287.]
- , 1963, "Actions, Reasons and Causes", *Journal of Philosophy*, vol. 60, no. 23, pp. 685-699. [Versión en castellano: "Acciones, razones y causas", en D. Davidson, *Ensayos sobre acciones y sucesos*, pp. 17-36.]
- Fodor, J.A., 1968, *Psychological Explanation*, Random House, Nueva York. [Versión en castellano: *La explicación psicológica*, trad. J.E. García Albea, Cátedra, Madrid, 1980.]
- Foster, L. y J.W. Swanson (comps.), 1970, *Experience and Theory*, University of Massachusetts Press, Amherst.
- O'Connor, J. (comp.), 1969, *Modern Materialism: Readings on Mind-Body Identity*, Harcourt, Brace and World, Nueva York.
- Putnam, H., 1970, "On Properties", en Rescher 1969, pp. 235-254. [Versión en castellano: *De las propiedades*, trad. M. Gorostiza, Instituto de Investigaciones Filosóficas-UNAM, México, 1983 (Cuadernos de Crítica, 19).]
- , 1967, "Psychological Predicates", en Capitan y Merrill 1967, pp. 37-48.
- , 1966, "The Mental Life of Some Machines", en O'Connor 1969, pp. 263-281. [Versión en castellano: *La vida mental de algunas máquinas*, trad. M. Gorostiza, Instituto de Investigaciones Filosóficas-UNAM, México, 1981 (Cuadernos de Crítica, 17).]

Putnam, H., 1965, "Brains and Behavior", en Butler 1965, pp. 1-19.

[Versión en castellano: *Cerebro y conducta*, trad. R. Amieva, Instituto de Investigaciones Filosóficas-UNAM, México, 1983 (Cuadernos de Crítica, 23).]

Rescher, N. (comp.), 1969, *Essays in Honor of C.G. Hempel*, Reidel, Dordrecht.