

# LA NATURALEZA DE LOS ESTADOS MENTALES<sup>1</sup>

Hilary Putnam

Los intereses típicos de los filósofos de la mente podrían expresarse mediante tres preguntas: (1) ¿Cómo sabemos que otras personas tienen dolores? (2) ¿Son los dolores estados cerebrales? (3) ¿Cuál es el análisis del concepto dolor? En este artículo no quiero discutir las preguntas (1) y (3). Diré algo acerca de la pregunta (2)<sup>2</sup>.

## I. PREGUNTAS SOBRE LA IDENTIDAD

“¿Es el dolor un estado cerebral?” (o bien: “La propiedad de tener un dolor en el tiempo *t*, ¿es un estado cerebral?”)<sup>3</sup>. Resulta imposible discutir sensatamente esta pregunta sin decir algo acerca de las peculiares reglas que han surgido en el curso del desarrollo de la “filosofía analítica” -reglas que, lejos de poner fin a todas las confusiones conceptuales, representan en sí mismas una considerable confusión conceptual. Estas reglas -que se hallan, por supuesto, implícitas más que explícitas en la práctica de la mayor parte de los filósofos analíticos- son: (I) que un enunciado de la forma “ser *A* es ser *B*” (v. gr. “tener dolor es estar en un determinado estado cerebral”) puede ser *correcto* sólo si se sigue, en algún sentido, del significado de los términos *A* y *B*; y (II) que un enunciado de la forma “ser *A* es ser *B*” puede ser filosóficamente *informativo* sólo si es en algún sentido reductivo (v. gr. “tener un dolor es tener una determinada sensación desagradable” no es filosóficamente informativo; “tener un dolor es tener una cierta disposición conductual” es filosóficamente informativo en el caso de ser verdadero). Estas reglas son excelentes si aún creemos que el programa de análisis reductivo (al estilo de los años treinta) puede llevarse a cabo; si no lo creemos, convierten entonces a la filosofía analítica en una farsa, por lo menos en la medida en que nos interesan las cuestiones acerca de “es”, esto es, acerca de la identidad.

En este trabajo usaré el término “propiedad” como un término irrestricto para cosas tales como tener dolor, hallarse en un estado cerebral particular, tener una disposición conductual particular, etcétera -esto es, para cosas que naturalmente pueden representarse mediante predicados o funtores monádicos o *n-arios*. Usaré el término “concepto” para cosas que pueden identificarse con clases de sinonimia de expresiones. Así, el concepto de *temperatura* puede identificarse (según sostengo) con la clase de sinonimia de la palabra “temperatura”<sup>4</sup>. (Esto es tanto como decir que el número 2 puede

---

<sup>1</sup> Traducción de Margarita Valdés.

<sup>2</sup> He discutido estos temas, y otros afines, en mis artículos “Brains and Behavior”, “Minds and Machines” y “The Mental Life of Some Machines” que aparecen en mi libro *Mind, Language and Reality*, Cambridge University Press., 1975.

[Hay traducción al español: véase *Mentes y máquinas* comp. A.R. Anderson UNAM, 1970; “La vida mental de algunas máquinas”, en: *Cuadernos de Crítica* No. 17, UNAM, 1971; “Cerebro y comportamiento” también en *Cuadernos de Crítica*, UNAM. (N. del T.)]

<sup>3</sup> En este trabajo quiero evitar la debatida cuestión acerca de la relación entre los dolores y los estados de dolor. Observo sólo de pasada que es fácil advertir la falacia de un argumento común en contra de la identificación de estos dos, a saber, que un dolor puede estar en el brazo de uno pero un estado (del organismo) no puede estar en el brazo de uno.

<sup>4</sup> Son bien conocidas ciertas observaciones de Alonzo Church sobre este asunto. Dichas observaciones no afectan (como podría suponerse en un principio) la identificación de conceptos con clases de sinonimia en cuanto tales, sino más bien apoyan la idea de que (en la semántica formal) es necesario conservar la distinción de Frege entre el uso normal y el uso “oblicuo” de las expresiones. Esto es, aun si decimos que el concepto de temperatura es la clase de sinonimia de la palabra “temperatura”, no por ello debemos caer en el error de

identificarse con la clase de todos los pares. Este es un enunciado bastante diferente del enunciado peculiar que 2 es la clase de todos los pares. Yo no sostengo que los conceptos *sean* clases de sinonimia, sea lo que fuere que esto pudiera significar, sino que pueden identificarse con clases de sinonimia para los propósitos de formalización del discurso relevante.)

La pregunta "¿qué es el concepto *temperatura*?" es una pregunta muy "chistosa". Podría considerarse que significa: "¿Qué es la temperatura? Por favor, considere que mi pregunta es conceptual." En ese caso, una respuesta podría ser (pretendamos que "calor" y "temperatura" son sinónimos) "la temperatura es el calor", o aun "el concepto de temperatura es el mismo que el concepto de calor" O uno podría considerar que significa: "¿Qué son realmente los conceptos? Por ejemplo, ¿qué es 'el concepto de temperatura'?" En este caso sólo Dios sabe lo que sería una "respuesta". (Tal vez sería el enunciado de que los conceptos *pueden identificarse con* clases de sinonimia.)

Desde luego que la pregunta "¿qué es la propiedad temperatura?" también es "chistosa", y una manera de interpretarla es considerarla como una pregunta acerca del concepto de temperatura. Pero no sería ésta la manera como la consideraría un físico.

El efecto de decir que la propiedad  $P_1$  puede ser idéntica a la propiedad  $P_2$  sólo si los términos  $P_1$  y  $P_2$  son, en algún sentido apropiado, "sinónimos", es, para cualquier propósito e intención, disolver las dos nociones de "propiedad" y "concepto" y fundirlas en una sola noción. La tesis de que los conceptos (intensiones) *son* lo mismo que las propiedades ha sido explícitamente defendida por Carnap (por ejemplo, en *Meaning and Necessity*). Esta parece ser una tesis desafortunada, ya que "la temperatura es la energía molecular cinética media" es evidentemente un ejemplo perfectamente bueno de un enunciado verdadero de identidad de propiedades, en tanto que "el concepto de temperatura es el mismo concepto que el concepto de energía molecular cinética media" es simplemente falso.

Muchos filósofos creen que el enunciado "el dolor es un estado cerebral" viola algunas reglas o normas del lenguaje. Sin embargo, los argumentos que se ofrecen son poco convincentes. Por ejemplo, si el hecho de que yo pueda saber que tengo dolor sin saber que me encuentro en el estado cerebral  $S$ , muestra que el dolor no puede ser el estado cerebral  $S$ , entonces, por el mismo argumento, el hecho de que yo pueda saber que la estufa está caliente sin que sepa que la energía cinética molecular es alta (o incluso que las moléculas existen) muestra que es *falso* que la temperatura sea la energía molecular cinética media, aun cuando la física diga lo contrario. De hecho, lo único que se sigue inmediatamente del hecho de que yo puedo saber que tengo dolor sin saber que estoy en el estado cerebral  $S$ , es que el concepto de dolor no es el mismo concepto que el concepto de estar en el estado cerebral  $S$ . Pero pudiera ser que el dolor, o el estado de tener dolor, o algún dolor, o algún estado doloroso fuese el estado cerebral  $S$ . Después de todo, el concepto de temperatura no es el mismo concepto que el concepto de energía cinética molecular media. Pero la temperatura es la energía cinética molecular media.

Algunos filósofos sostienen que tanto "el dolor es un estado cerebral" como "los estados de dolor son estados cerebrales" son ininteligibles. La respuesta es explicar a estos filósofos, tan bien como podamos, dada la vaguedad de toda metodología científica, qué tipos de consideraciones nos conducen a hacer una reducción empírica (esto es, a decir cosas tales como "el agua es  $H_2O$ ", "la luz es una radiación electromagnética", "la temperatura es la energía molecular cinética media"). Si, a pesar de estos ejemplos, alguien continúa manteniendo, sin dar razones, que no puede imaginarse

---

suponer que "el concepto de temperatura" es sinónimo de "la clase de sinonimia de la palabra 'temperatura' ", ya que entonces "el concepto de temperatura" y "der Begriff der Temperatur" no serían sinónimos, cosa que sí son. Más bien debemos decir que el concepto de "temperatura" se refiere a la clase de sinonimia de la palabra "temperatura" (en esta reconstrucción particular); pero esa clase se identifica, no como "la clase de sinonimia a la que tal y cual palabra pertenece" sino de otra manera (v. gr como la clase de sinonimia cuyos miembros tienen tal y cual uso característico).

circunstancias paralelas para el uso de "los dolores son estados cerebrales" (o, tal vez, "los estados de dolor son estados cerebrales"), habrá motivos para considerarlo perverso.

Algunos filósofos mantienen que " $P_1$  es  $P_2$ " es algo que puede ser verdadero, cuando el "es" contenido en esa expresión es el "es" de una reducción empírica, sólo en el caso de que las propiedades  $P_1$  y  $P_2$  estén (a) asociadas con una región espaciotemporal y (b) la región sea una y la misma en ambos casos. Así, "la temperatura es la energía molecular cinética media" es una reducción empírica admisible, ya que la temperatura y la energía molecular se asocian a la misma región espaciotemporal, pero "tener un dolor en mi brazo es estar en un estado cerebral" no lo es, ya que las regiones espaciales aludidas son diferentes.

Este argumento no parece muy sólido. ¡Ciertamente nadie se detendrá a pensar, antes de decir que las imágenes en los espejos son la luz reflejada por un objeto y luego por la superficie del espejo, en el hecho de que una imagen puede "localizarse" un metro *detrás* del espejo! (Además, uno puede encontrar siempre *alguna* propiedad común de las reducciones que uno está dispuesto a admitir –v. gr. la temperatura es la energía molecular cinética media–, que no sea una propiedad de alguna identificación que uno desea desautorizar. Esto no resulta demasiado impresionante a menos que uno tenga un argumento para mostrar que los propósitos mismos de semejante identificación dependen de la propiedad común en cuestión.)

Otros filósofos más han sostenido que todas las predicciones derivables de la conjunción de leyes neurofisiológicas con enunciados tales como "los estados de dolor son estados cerebrales", pueden derivarse igualmente a partir de la conjunción de las mismas leyes neurofisiológicas con "tener dolor está correlacionado con tal y cual estado cerebral" y que, por ende (¡*sic!*), no puede haber fundamentos metodológicos para decir que los dolores, o los estados de dolor, *son* estados cerebrales, por oposición a decir que están *correlacionados* (invariablemente) con estados cerebrales. Este argumento mostraría también que la luz está solamente correlacionada con la radiación electromagnética. El error está en ignorar el hecho de que, aunque las teorías en cuestión pueden efectivamente conducir a las mismas predicciones, excluyen o dejan abiertas diferentes *preguntas*. "La luz está correlacionada invariablemente con la radiación electromagnética" dejaría abiertas las preguntas "¿qué es entonces la luz, si no es lo mismo que la radiación electromagnética?" y "¿qué es lo que hace que la luz acompañe a la radiación electromagnética?", las cuales se excluyen al decir que la luz es la radiación electromagnética. De manera similar, el propósito de decir que los dolores son estados cerebrales es precisamente excluir del campo de lo empíricamente significativo las preguntas "¿qué es entonces el dolor si no es lo mismo que el estado cerebral?" y "¿qué es lo que hace que el dolor acompañe al estado cerebral?" Si hay bases que sugieran que estas preguntas representan, por así decirlo, la manera equivocada de considerar el asunto, entonces esas bases son las bases para una identificación teórica de los dolores con los estados cerebrales.

Si ninguno de los argumentos para establecer lo contrario resulta convincente, ¿habremos entonces de concluir que es significativo (y tal vez verdadero) decir que los dolores son estados cerebrales, o que los estados de dolor son estados cerebrales?

(1) Es perfectamente significativo (no viola ninguna "regla del lenguaje", no supone ninguna "extensión del uso común") decir "los dolores son estados cerebrales".

(2) No es significativo (supone un "cambio de significado" o "una extensión del uso", etcétera) decir "los dolores son estados cerebrales".

Mi posición no se expresa ni en (1) ni en (2). Me parece que las nociones "cambio de significado" y "extensión del uso" están tan mal definidas que de hecho uno no puede decir *ni* (1) *ni* (2). No veo ninguna razón para creer que el lingüista, o el hombre de la calle, o el filósofo, posean hoy día una noción de "cambio de significado" aplicable a casos tales como los que hemos estado discutiendo.

La faena para la cual se desarrolló, en la historia del lenguaje, la noción de cambio de significado, era una faena *mucho* más burda que ésta.

Pero, si no aseveramos ni (1) ni (2) –en otras palabras, si en este caso consideramos el asunto del "cambio de significado" como un pseudoasunto-, ¿cómo hemos de discutir la cuestión con la que empezamos? "¿Es el dolor un estado cerebral?"

La respuesta consiste en permitir enunciados de la forma "el dolor es A", en donde "dolor" y "A" no son de ninguna manera sinónimos, y ver si puede encontrarse algún enunciado de dicha índole que pudiera ser aceptable sobre bases empíricas y metodológicas. En lo que sigue procederemos a hacer esto.

## II. ¿ES EL DOLOR UN ESTADO CEREBRAL?

Discutiremos, pues, la pregunta "¿es el dolor un estado cerebral?" Hemos acordado hacer a un lado el asunto relativo al "cambio de significado".

Puesto que no estoy discutiendo a qué equivale el concepto de dolor, sino qué es el dolor, en un sentido de "es" que exige la construcción empírica de una teoría (o, al menos, la especulación empírica), no ofreceré disculpas por proponer una hipótesis empírica. De hecho, mi estrategia será alegar que el dolor *no* es un estado cerebral, no sobre bases *a priori*, sino sobre la base de que otra hipótesis es más plausible. El desarrollo y la verificación detallada de mi hipótesis sería una tarea tan utópica como el desarrollo y la verificación detallada de la hipótesis de que el dolor es un estado cerebral. Sin embargo, la postulación, no de hipótesis detalladas y científicamente "acabadas", sino de esquemas para hipótesis, ha sido por mucho tiempo una función de la filosofía. Argumentaré, en breve, que el dolor no es un estado cerebral, en el sentido de un estado fisicoquímico del cerebro (o aun de todo el sistema nervioso), sino *otra* clase enteramente diferente de estado. Propongo la hipótesis de que el dolor, o el estado de tener dolor, es un estado funcional de todo un organismo.

Para explicar esto es necesario introducir algunas nociones técnicas. En trabajos anteriores he explicado la noción de una Máquina de Turing y he discutido el uso de esta noción como modelo para un organismo. La noción de un Autómata Probabilista se define de manera similar a la de una Máquina de Turing, excepto en que se permite que las transiciones entre "estados" se den con varias probabilidades en lugar de ser "deterministas". (Por supuesto, una Máquina de Turing es sencillamente una clase especial de Autómata Probabilista, esto es, una máquina cuyas probabilidades de transición son 0,1.) Asumiré que la noción de Autómata Probabilista se ha generalizado de tal manera que permite "entradas sensoriales" y "salidas motoras" –es decir, la Tabla de la Máquina específica, para cada combinación posible de un "estado" con un conjunto completo de "entradas sensoriales", una "instrucción" que determina la probabilidad del "estado" siguiente, así como la probabilidad de las "salidas motoras". (Esto reemplaza la idea de la Máquina que escribe en una cinta.) También asumiré que la realización física de los órganos sensoriales responsables de las diferentes entradas, así como la de los órganos motores, se halla especificada, pero que los "estados" y las "entradas" mismas se especifican, como es costumbre, sólo implícitamente, esto es, mediante el conjunto de probabilidades de transición dado por la Tabla de Máquina.

Puesto que un sistema empíricamente dado puede ser a la vez una "realización física" de muchos Autómatas Probabilistas diferentes, introduzco la noción de *Descripción* de un sistema. Una descripción de  $S$ , en donde  $S$  es un sistema, es cualquier enunciado verdadero que dice que  $S$  posee los estados diferentes  $S_1, S_2, \dots, S_n$ , los cuales están relacionados entre sí, y con las salidas motoras y las entradas sensoriales, mediante las probabilidades de transición dadas en tal y cual Tabla de Máquina. La Tabla de Máquina mencionada en la descripción se llamará la Organización Funcional de  $S$  relativa a esa Descripción, y el  $S$ , tal que  $S$  se halla en el estado  $S_i$  en un momento dado, se llamará el Estado Total de  $S$  (en el momento) relativo a esa Descripción. Hay que tener en cuenta que conocer el Estado Total de un sistema relativo a una Descripción, supone conocer en buena

medida cómo es probable que el sistema se "comporte" dadas varias combinaciones de entradas sensoriales, pero no supone conocer la realización física de los  $S_i$  como, por ejemplo, estados fisicoquímicos del cerebro: Los  $S_i$ , repitamos, se especifican sólo *implícitamente* mediante la Descripción, esto es, se especifican sólo mediante el conjunto de probabilidades de transición dado en la Tabla de Máquina.

La hipótesis de que "ser un dolor es ser un estado funcional de un organismo" puede ahora explicarse de manera más exacta de la siguiente forma:

- (1) Todos los organismos capaces de sentir dolor son Automatas Probabilistas.
- (2) Todo organismo capaz de sentir dolor posee por lo menos una Descripción de un cierto tipo (esto es, ser capaz de sentir dolor es *poseer* una clase especial de Organización Funcional).
- (3) Ningún organismo capaz de sentir dolor puede descomponerse en partes que posean separadamente Descripciones del tipo de las mencionadas en (2).
- (4) Para cada Descripción de la clase mencionada en (2), existe un subconjunto de entradas sensoriales tal que un organismo con dicha Descripción siente dolor cuando, y sólo cuando, algunas de sus entradas sensoriales están en ese subconjunto.

Esta hipótesis es, sin lugar a dudas, vaga, aunque con toda seguridad no es más vaga que la hipótesis del estado cerebral tal como se halla en el presente. Por ejemplo, a uno le gustaría saber más acerca de la clase de Organización Funcional que un organismo ha de tener para ser capaz de sentir dolor, y más sobre los rasgos que distinguen el subconjunto de las entradas sensoriales a las que nos referimos en (4). Con respecto a la primera cuestión probablemente podamos decir que la Organización Funcional debe incluir algo semejante a una "función de preferencia", o por lo menos a un ordenamiento parcial de preferencia, y algo semejante a una "lógica inductiva" (esto es, la Máquina tiene que ser capaz de "aprender de la experiencia")<sup>5</sup>. Parece necesario, además, exigir que las máquinas posean "sensores de dolor", esto es, órganos sensoriales que normalmente indiquen daños al cuerpo de la Máquina, o temperaturas, presiones, etcétera, que representen un peligro, y que dichos sensores transmitan un subconjunto especial de las entradas, el subconjunto al que nos referimos en (4). Finalmente, y con respecto a la segunda cuestión, quisiéramos exigir por lo menos que las entradas en el subconjunto distinguido tengan un valor muy bajo en la función o el ordenamiento preferencial de la máquina<sup>6</sup>. El propósito de la condición (3) es el de eliminar "organismos" (si así pueden llamarse) tales como un enjambre de abejas, como sujetos singulares de dolor. La condición (1) es obviamente redundante y se introduce sólo por razones de la exposición. (De hecho, es una condición vacía, ya que todo es un Automata Probabilista bajo *alguna* Descripción.)

Sostengo, de pasada, que esta hipótesis, a pesar de ser vaga, es mucho menos vaga que la hipótesis del "estado fisicoquímico" tal como se presenta hoy día, y mucho más susceptible a la investigación tanto matemática como empírica. En efecto, investigar esta hipótesis consiste en tratar de producir modelos "mecánicos" de organismos y, después de todo, ¿no es esto acaso lo que persigue la psicología? El paso difícil, por supuesto, consistirá en pasar de los modelos para organismos *específicos* a una *forma normal* para la descripción psicológica de los organismos, ya que esto es lo que se requiere para hacer más precisas las condiciones (2) y (4). Sin embargo, esto también parece ser una parte inevitable del programa de la psicología.

Ahora haré una comparación de la hipótesis que acabo de presentar con: (i) la hipótesis de que el dolor es un estado cerebral y (ii) con la hipótesis de que el dolor es una disposición conductual.

---

<sup>5</sup> El significado de estas condiciones para los modelos de autómatas se discute en mi artículo "The Mental Life of Some Machines".

<sup>6</sup> Hay otras condiciones que se discuten en "The Mental Life of Some Machines".

### III. ESTADO FUNCIONAL VS. ESTADO CEREBRAL

Acaso sea posible preguntar si no soy un tanto injusto al considerar que el teórico que identifica los estados mentales con estados cerebrales habla acerca de estados *fisicoquímicos* del cerebro. Sin embargo, (a) esta es la única clase de estados mencionados siempre por los teóricos de los estados cerebrales. (b) Estos teóricos generalmente mencionan (con cierto orgullo que no deja de recordar el del "ateo del pueblo") la incompatibilidad de su hipótesis con cualquier forma de dualismo o mentalismo. Esto resulta natural si son los estados fisicoquímicos del cerebro los que se consideran. Sin embargo, los estados funcionales de sistemas completos son algo bastante diferente. En particular, la hipótesis funcionalista no es incompatible con el dualismo. Aunque resulta obvio que la hipótesis es de inspiración "mecanista", es un hecho notable que un sistema consistente en un cuerpo y un "alma", si es que hay tales cosas, puede perfectamente ser un Automata Probabilista. (c) Un argumento presentado por Smart es que la teoría de los estados cerebrales asume solamente propiedades "físicas", y a Smart le parecen ininteligibles las propiedades "no físicas". Los Estados Totales y las "entradas" definidos anteriormente no son, por supuesto, ni mentales ni físicos *per se*, y no puedo imaginar que ningún funcionalista proponga este argumento. (d) Si el teórico de los estados mentales quiere referirse a estados distintos a los fisicoquímicos (o al menos permitidos), entonces su hipótesis es completamente vacía, a menos que especifique a *qué* clase de "estados" *quiere* referirse.

Considerando de esta manera la hipótesis de los estados cerebrales, ¿qué razones tenemos entonces para preferir la hipótesis de los estados funcionales a la hipótesis de los estados cerebrales? Consideremos lo que el teórico de los estados cerebrales tiene que hacer para dar validez a sus afirmaciones. Tiene que especificar un estado fisicoquímico tal que *cualquier* organismo (no solamente un mamífero) siente dolor si y sólo si (a) posee un cerebro con una estructura fisicoquímica adecuada y (b) su cerebro está en ese estado fisicoquímico. Esto significa que el estado fisicoquímico en cuestión tiene que ser un estado posible de un cerebro de mamífero, de un cerebro de reptil, de un cerebro de molusco (los pulpos son moluscos y ciertamente sienten dolor), etcétera. Al mismo tiempo, tiene que ser un estado que *no* sea posible (físicamente posible) para el cerebro de ninguna creatura físicamente posible que no pueda sentir dolor. Aun cuando pueda encontrarse un estado semejante, tiene que ser nomológicamente cierto que será también un estado del cerebro de cualquier vida extraterrestre que pudiera descubrirse y que fuese capaz de sentir dolor, antes de que podamos siquiera mantener la suposición de que este estado *sea* el dolor.

No es completamente imposible que se encuentre un estado semejante. Aun cuando los pulpos y los mamíferos constituyen ejemplos de una evolución paralela (más que secuencial), por ejemplo, estructuras virtualmente idénticas (hablando desde una perspectiva física) han evolucionado en el ojo del pulpo y en el ojo de los mamíferos, pese al hecho de que en los dos casos dicho órgano ha evolucionado a partir de diferentes clases de células. Así, pues, es por lo menos posible que las evoluciones paralelas, a través de todo el universo, pudieran *siempre* conducir a *uno y el mismo* "correlato" físico del dolor. Sin embargo, esta es ciertamente una hipótesis ambiciosa.

Finalmente, la hipótesis se torna aun más ambiciosa cuando caemos en la cuenta de que el teórico de los estados cerebrales no solamente dice que el *dolor* es un estado cerebral; lo que le interesa mantener, por supuesto, es que *todo* estado psicológico es un estado cerebral. Entonces, si podemos encontrar aunque sea un solo predicado psicológico que pueda aplicarse claramente tanto al mamífero como al pulpo (digamos, "hambriento"), pero cuyo "correlato" fisicoquímico sea diferente en los dos casos, la teoría del estado cerebral se derrumba. Me parece sumamente probable que podamos hacer eso. Concedo que en tal caso el teórico de los estados cerebrales puede salvarse mediante supuestos *ad hoc* (por ejemplo, definiendo la disyunción de dos estados como un solo "estado fisicoquímico"), pero esto no tiene que tomarse en serio.

Pasando ahora a las consideraciones *en favor* de la teoría del estado funcional, comencemos por el hecho de que identificamos el dolor, el hambre, la cólera, el calor, etcétera, de los organismos sobre la base de su *conducta*. Pero es una verdad trillada que las similitudes en la conducta de dos sistemas constituyen por lo menos una razón para sospechar similitudes en la organización funcional de los dos sistemas, y una razón mucho *más débil* para sospechar similitudes en los detalles físicos que de hecho tengan. Más aun, suponemos que los distintos estados psicológicos – por lo menos los básicos, tales como el hambre, la sed, la agresión, etcétera- tienen "probabilidades de transición" más o menos semejantes (desde luego, dentro de límites amplios y mal definidos), entre sí y con la conducta en el caso de especies diferentes, porque esto es producto de la manera como identificamos estos estados. Así, no diríamos que un animal está *sediento* si su conducta "insaciada" no pareciera estar dirigida a beber y no fuese seguida de una "saciedad de líquido". Así, cualquier animal que consideremos capaz de tener estos estados diferentes por lo menos *parecerá* tener una cierta clase de organización funcional. Por otro lado, como señalamos anteriormente, si el programa de descubrir leyes psicológicas que no sean específicas para una especie –esto es, de hallar una forma normal para teorías psicológicas sobre especies diferentes- llegara a tener éxito en algún momento, traería como consecuencia un delineamiento de la clase de organización funcional que es necesaria y suficiente para un estado psicológico dado, así como una definición precisa de la noción de "estado psicológico". En contraste, el teórico de los estados cerebrales tiene que esperar que algún día se desarrollen leyes neurofisiológicas que sean independientes de la especie, lo cual parece mucho menos razonable que tener la esperanza de que las leyes psicológicas (de una clase lo suficientemente general) puedan ser independientes de la especie o, dicho de una forma aun más débil, que pueda encontrarse una *forma* de escribir leyes psicológicas que sea independiente de la especie.

#### IV. ESTADO FUNCIONAL vs. DISPOSICIÓN DE CONDUCTA

La teoría que afirma que tener un dolor no es un estado cerebral ni un estado funcional, sino una disposición conductual, tiene una ventaja aparente: parece concordar con la manera como verificamos que los organismos tienen dolor. En la práctica no sabemos nada acerca del estado cerebral de un animal cuando decimos que siente dolor, y tenemos poco o ningún conocimiento acerca de su organización funcional, excepto de una manera intuitiva burda. De hecho, sin embargo, esta "ventaja" no constituye ninguna ventaja, ya que, aunque los enunciados acerca de cómo verificamos que  $x$  es  $A$  pueden tener mucho que ver con lo que es el concepto de ser  $A$ , tienen muy poco que ver con lo que es la propiedad  $A$ . Argumentar, sobre la base recién mencionada, que el dolor no es ni un estado cerebral ni un estado funcional, es semejante a argumentar que el calor no es la energía molecular cinética media a partir del hecho de que la gente común no descubre (según piensa) la energía molecular cinética media de algo cuando verifica que está caliente o frío. No es necesario que lo haga: lo necesario es que los signos que toma como indicaciones de calor sean de hecho explicados por la energía molecular cinética media. Así, de manera semejante, nuestra hipótesis requiere que los signos que se toman como indicaciones conductuales del dolor se expliquen por el hecho de que el organismo está en un estado funcional de la clase apropiada, pero no que los hablantes sepan que esto es así.

Las dificultades propias de las explicaciones en términos de "disposiciones conductuales" son tan ampliamente conocidas que aquí me limitaré a recordarlas. La dificultad (de hecho parece ser más que una simple "dificultad") de especificar la disposición conductual requerida excepto como "la disposición de  $X$  a comportarse como si  $X$  sintiera *dolor*", es desde luego la más importante. En contraste, *podemos* especificar, al menos en términos generales, el estado funcional con el que pretendemos identificar el dolor, sin usar la noción de dolor. El estado funcional que tenemos en mente es, a saber, el estado de recibir entradas sensoriales que juegan determinado papel en la Organización Funcional del organismo. Este papel se caracteriza, al menos parcialmente, por el

hecho de que los órganos sensoriales responsables de las entradas en cuestión son órganos cuya función es la de detectar daño al cuerpo, o extremos peligrosos de temperatura, presión, etcétera, y por el hecho de que, cualquiera que sea la realización física de las "entradas" mismas, el organismo les asigna un alto valor negativo. Como he subrayado en "The Mental Life of Some Machines", esto *no* significa que la máquina habrá siempre de *evitar* hallarse en la condición en cuestión ("dolor"); sólo significa que dicha condición será evitada a menos que el no evitarla sea necesario para alcanzar alguna otra meta más altamente valorada. Dado que la conducta de la máquina (en este caso un organismo) dependerá no solamente de las entradas sensoriales sino también del Estado Total (es decir, de otros valores, creencias, etcétera), parece inútil hacer cualquier enunciado general acerca de cómo tiene que comportarse un organismo en tal condición, pero esto no significa que tengamos que abandonar toda esperanza de caracterizar la condición. De hecho, acabamos de caracterizarla<sup>7</sup>.

La teoría de la disposición conductual no sólo parece irremediabilmente vaga, sino que parece claramente falsa si la "conducta" a la que se refiere es la conducta periférica y los estímulos relevantes son los estímulos periféricos (*v. gr.* no se dice nada acerca de lo que el organismo haría si se le opera el cerebro). Por ejemplo, dos animales con todos los nervios motores cortados tendrán la misma conducta presente y potencial (esto es, ninguna, por así decirlo); pero si uno de ellos tiene cortadas las fibras del dolor y el otro las tiene intactas, entonces uno de ellos sentirá dolor y el otro no. De la misma manera, si una persona tiene cortadas las fibras del dolor y otra suprime deliberadamente, a causa de una fuerte compulsión, todas las respuestas al dolor, entonces la conducta periférica presente y potencial puede ser la misma, pero una sentirá dolor y la otra no. (Algunos filósofos sostienen que este último caso es conceptualmente imposible, pero la única prueba de esto parece ser que ellos no pueden, o no quieren, concebirlo.)<sup>8</sup> Si en lugar del dolor tomamos alguna sensación cuya "expresión corporal" sea más fácil de suprimir –digamos, la sensación de frío en el dedo meñique izquierdo de una persona– el argumento se vuelve aun más claro.

Finalmente, aun si *hubiese* alguna disposición conductual invariablemente correlacionada con el dolor (¡independientemente de la especie!), y susceptible de especificarse sin usar el término "dolor", seguiría siendo más plausible identificar el tener un dolor con algún estado cuya presencia explicase esta disposición conductual –el estado cerebral o el estado funcional– que con la disposición conductual misma. Tales consideraciones con respecto a la plausibilidad pueden ser de alguna manera subjetivas, pero si todas las circunstancias fuesen iguales (por supuesto, no lo son), ¿por qué no habríamos de permitir que las consideraciones con respecto a la plausibilidad jugasen el papel decisivo?

## V. CONSIDERACIONES METODOLÓGICAS

Hasta aquí hemos considerado solamente lo que podría llamarse las razones "empíricas" para decir que tener dolor es un estado funcional, más que un estado cerebral o una disposición conductual; es decir, que parece más probable que el estado funcional que hemos descrito esté invariablemente "correlacionado" con el dolor, independientemente de la especie, que el que haya un estado

---

<sup>7</sup> En "The Mental Life of Some Machines" se discute una característica más (de alguna manera independiente) de las entradas de dolor en términos de modelos de autómatas: a saber, la espontaneidad de la inclinación a retraer la parte lastimada, etcétera. Esto hace surgir el problema de dar un análisis funcional de la noción de inclinación espontánea, lo cual se discute en dicho artículo. Desde luego, otras características más acuden fácilmente a la mente, por ejemplo, que las sensaciones de dolor están (o parecen estar) localizadas en las diferentes partes del cuerpo.

<sup>8</sup> Cfr. la discusión de 108 "super-espartanos" en mi artículo "Brains and Behavior", en mi libro citado en la nota 1.



fisicoquímico del cerebro (¿tiene que tener *cerebro* un organismo para sentir dolor? Quizá le basten algunos ganglios) o una disposición conductual correlacionados de esa manera. Si esto es correcto, entonces se sigue que la identificación que hemos propuesto es por lo menos un candidato digno de considerarse. ¿Qué pasa con las consideraciones metodológicas?

Las consideraciones metodológicas son, a grandes rasgos, semejantes en todos los casos de reducción, de manera que no debemos esperar sorpresas aquí. Primero, la identificación de estados psicológicos con estados funcionales significa que las leyes de la psicología pueden derivarse de enunciados de la forma "tales y cuales organismos tienen tales y cuales Descripciones" junto con los enunciados de identificación ("tener dolor es tal y cual estado funcional", etcétera). Segundo, la presencia del estado funcional (esto es, de entradas que juegan el papel que hemos descrito en la Organización Funcional del organismo) no está meramente "correlacionado con" la conducta de dolor por parte del organismo, sino que de hecho la explica. Tercero, la identificación sirve para excluir preguntas que representan (suponiendo que el punto de vista naturalista es correcto) una manera totalmente equivocada de considerar la cuestión, esto es, preguntas tales como "¿qué es el dolor si no es el estado cerebral ni el estado funcional?" y "¿qué es lo que causa que el dolor vaya siempre acompañado de este tipo de estado funcional?" En suma, la identificación ha de aceptarse tentativamente como una teoría que conduce tanto a predicciones fructíferas como a *preguntas* fructíferas, y que sirve para desalentar preguntas empíricamente sin sentido, en donde por "empíricamente sin sentido" quiero decir "sin sentido" no solamente desde el punto de vista de la verificación, sino desde el punto de vista de lo que de hecho es.