

¿POR QUÉ LOS ESTADOS MENTALES NO SON CLASES NATURALES?

OLBETH HANSBERG

INSTITUTO DE INVESTIGACIONES FILOSÓFICAS
UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

El que existan o no leyes psicofísicas tiene consecuencias para el problema mente-cuerpo en general y en particular para las diversas teorías de identidad de lo mental y lo físico, sobre todo en relación con la posibilidad de reducción de la psicología a alguna ciencia física.

Si las leyes psicofísicas no existen, y queremos seguir sosteniendo alguna teoría fiscalista, tendríamos que renunciar a una teoría de identidad de tipos y a un fiscalismo reduccionista, en favor de, por ejemplo, una teoría de identidades particulares unida a una versión del fiscalismo que niegue la posibilidad de reducción y de esta manera afirme la autonomía de las explicaciones psicológicas.¹

En este artículo me propongo examinar uno de los argumentos en contra de la existencia de leyes psicofísicas que se basa en la teoría de las clases naturales propuesta por Putnam y Kripke. El esquema general del argumento² es el siguiente:

- 1) Cualquier término que aparezca en una ley deberá ser un término de clase natural.
- 2) Los términos mentales no son términos de clases naturales.

Por lo tanto,

- 3) Ninguna ley puede incluir términos mentales.

¹ Por ejemplo el monismo anómalo de Davidson en *Essays on Action and Events*, Oxford University Press, 1980. Ver especialmente ensayos 11, 12 y 13. Traducción al español del artículo "Mental Events" en *Cuadernos de Crítica* No. 11

² Este argumento lo propuso Mc Ginn en "Mental States, Natural Kinds, and Psychophysical Laws", *Proceedings of the Aristotelian Society*, 1978. Por ley ha de entenderse aquí una ley estricta del tipo de las de las ciencias naturales.

Este arguemo excluye la posibilidad de leyes que contengan términos mentales, esto es, va en contra tanto de leyes psicofísicas como de leyes puramente psicológicas.

A continuación argumentaré a favor de las premisas de este argumento, sobre todo a favor de la segunda, que es la que más me interesa en este artículo. Sin embargo, antes de exponer y analizar este argumento con mayor detalle, es necesario que me refiera brevemente a una teoría de las clases naturales y de los términos con los que nos referimos a ellas.³

I

La tesis de Putnam acerca de cómo se introducen estas palabras en el lenguaje es la siguiente: empezamos con un conjunto de buenos ejemplos o "paradigmas" de alguna substancia u objeto como, por ejemplo, agua, oro o tigre y los "bautizamos" como 'agua', 'oro', 'tigre'. Después usamos esos términos para referirnos a cualquier cosa que tenga la misma naturaleza o sea del *mismo tipo* que los ejemplos paradigmáticos. La referencia a "mismo tipo" se elimina mediante la investigación científica de los miembros de ese conjunto paradigmático, es decir, la ciencia empírica investiga las estructuras explicativas de los paradigmas, su esencia real subyacente, que puede consistir, por ejemplo, en la estructura atómica, genética, etc. Así, algo es agua si y sólo si tiene la misma estructura atómica que los ejemplos paradigmáticos, a saber, si y sólo si es H_2O . Ahora bien, estos paradigmas son ejemplos de un tipo de objetos que reunimos porque tienen ciertas características superficiales en común, sin embargo, el hecho de que los usemos como paradigmas no nos obliga a incluirlos después necesariamente en la extensión del término. Un pedazo de metal, por ejemplo, que todos llamamos 'oro' y que usamos para enseñarle a alguien la palabra 'oro' podría resultar que no fuera oro, si después descubrimos que su estructura atómica no es como la de los otros ejemplos paradigmáticos. En ese caso diríamos que hemos usado equivocadamente el término 'oro' y que ese pedazo de metal en realidad no era oro. Para que un objeto pueda considerarse realmente como miembro de una clase natural, es necesario que tengamos una ciencia más o menos avanzada, pues es ésta la que descubre la esencia real que a su vez determina la extensión de un término de clase natural y el criterio de pertenencia a ella.

En la medida en que la ciencia precisa en qué consiste la forma explicativa de una clase natural dada, se abre la posibilidad de una reclasificación

³ Ver H. Putnam, *Mind Language and Reality*, Philosophical Papers, Vol. 2, Cambridge University Press, 1975 y Saul Kripke, "Naming and Necessity" en *Semantics of Natural Language*, Davidson & Harman (eds.), Boston, 1972. Traducción al español: *El nombrar y la necesidad*, UNAM, 1985. También M. Platts, "Kind Words and Understanding", *Crítica* XII, núm. 36, 1980 y "Explanatory Kinds" en *Brit. J. Phil. Sci.*, núm. 34, 1983.

científica, esto es, que la extensión de un término de clase natural dado incluya objetos que previamente se habían descartado y que excluya otros que, aunque tengan las mismas características superficiales que los otros miembros de la clase, no compartan su misma esencia subyacente.

Podemos distinguir algunos rasgos lógico-semánticos característicos de este tipo de palabras:

1. Debido a que la investigación científica nos permite desubrir las formas explicativas básicas de los miembros de una clase natural, las definiciones operativas de palabras de clases naturales que proporcionen condiciones necesarias y suficientes, deberán ser *a posteriori*.
2. De lo anterior se sigue también que la referencia de este tipo de términos se fija parcialmente *antes* de que conozcamos las condiciones necesarias y suficientes para su aplicación —esto es, fijamos en parte la extensión al apuntar a los miembros del conjunto paradigmático— y muchas veces fijamos esta extensión sin saber en qué consiste, o teniendo ideas equivocadas acerca de la esencia real de los objetos que forman esa clase.
3. El hecho de introducir estos términos mediante casos paradigmáticos les confiere un elemento “déictico (*indexical*)...” “‘agua’ es la substancia que tiene una cierta relación de similitud con el agua que hay *por aquí*. El agua en otro tiempo o lugar o aun en otro mundo posible debe, para ser agua, tener la relación mismo [mismo líquido que] con nuestra ‘agua’”⁴
4. Esto se relaciona, a su vez, con la idea de que las palabras de clases naturales son designadores rígidos en el sentido de Kripke.⁵ Para Kripke, un designador rígido es un nombre o descripción que se refiere al mismo objeto o tipo de objetos en todo mundo posible en el que ese objeto o ese tipo de objetos existe. Así, si ‘oro’ es un designador rígido, se refiere siempre a la misma substancia, independientemente de cuales sean sus propiedades fenoménicas superficiales. No podemos definir lo que sea oro mediante una lista de propiedades que se especifican analíticamente⁶ puesto que, cualquiera otra substancia que tenga esas mismas propiedades superficiales, no será oro a menos que sea del mismo tipo de substancia que aquella que hemos designado rígidamente con la palabra ‘oro’, a saber, el elemento con el número atómico 79.

⁴ Putnam, *op. cit.*, p. 234.

⁵ Kripke, *op. cit.*, p. 270.

⁶ Según Kripke se especifican *a priori*, lo que para él quiere decir que alguien entiende una palabra sólo si conoce su definición, a saber, la lista de propiedades que se supone son condiciones necesarias y suficientes para la aplicación de la palabra. Aquí *a priori* es más cercano al sentido tradicional de *analítico*.

Sin embargo, estos rasgos lógico-semánticos que hemos mencionado, sólo nos permiten distinguir, frente a otro tipo de palabras, cuáles son los términos de un lenguaje que funcionan como término de clase natural, pero todavía no podemos responder a la pregunta de cuándo una clase de objetos constituye realmente una clase natural. Para que una palabra sea una palabra de clase natural, es necesario cuando menos que la palabra sea usada como un término de clase natural y que su extensión sea efectivamente una clase natural. Una de las cosas que debemos enfatizar es que existen objetos en el mundo que forman clases naturales y que estas clases son reales, es decir, existen independientemente de que nosotros lo sepamos o no, e independientemente de que tengamos una teoría que explique en qué consiste su esencia real. Estas teorías y explicaciones podrán modificarse de tal manera que nuestra descripción del mundo sea cada día más adecuada. Así, podría resultar que una clase de objetos que en un momento dado creíamos que era una clase natural, en realidad no lo sea, y que objetos que no parecían tener rasgos comunes, en realidad tengan una misma estructura explicativa.

¿Cuál es entonces la diferencia entre una clase natural y una clase “artificial”? Para contestar esta pregunta, tenemos que recurrir a una serie de consideraciones ontológicas y epistemológicas. “Una clase de objetos ϕ forma una clase natural cuando ‘ ϕ s’ es la respuesta máximamente explicativa a la pregunta ‘¿qué son estos?’. . . La clase de los tigres (electrones) forma una clase natural de tigres (electrones) porque si preguntamos con respecto a los miembros de la clase ‘¿qué son estos?’, la respuesta ‘tigres’ (‘electrones’) es la respuesta máximamente explicativa: es la respuesta que, al referirse a nuestra teoría de tigres (teoría de electrones), puede servir máximamente para explicar qué son y cómo se comportan esos objetos.”⁷ En cambio, si preguntamos acerca de objetos como lápices, botellas, ‘¿qué son estos?’, la respuesta ‘lápices’, ‘botellas’, no es suficientemente explicativa puesto que se dejan de explicar muchas propiedades de esos objetos, porque prácticamente no tenemos una teoría acerca de lápices o botellas que explique cómo es que tienen las características superficiales que de hecho tienen. Parece ser, entonces, que las clases artificiales no admiten una caracterización común máximamente explicativa.

Otra noción que debemos explicar es la del nivel de una clase natural. Los limones, las naranjas y las toronjas forman clases naturales pero pertenecen también, junto con otros objetos, a una clase natural de un nivel superior, la de las frutas cítricas; ésta a su vez forma parte de otra clase natural de un nivel todavía superior, la de las frutas, etc.

Ahora bien, una clase de objetos es una clase natural cuando los miembros de esa clase tienen una forma explicativa común. La clase de las frutas cítricas, por ejemplo, es una clase natural debido a que sus miembros tienen

⁷ Platts, “What is a Natural Kind Word”.

ciertos rasgos explicativos comunes. Pero estos miembros pertenecen a su vez a otras clases naturales, la de los limones, la de las naranjas, etc., de tal manera que aparte de las formas explicativas que comparten *qua* frutas cítricas, deberán tener otros rasgos explicativos distintos que son específicos sólo de las naranjas o sólo de los limones, etc. Aquí podría servirnos la noción de peso explicativo adicional,⁸ que se refiere al aumento en fuerza explicativa que se obtiene al añadirle algo a una descripción, de tal manera que (en general) se restrinja la extensión de la descripción. Así, por ejemplo, si tenemos una clase natural especificada por 'felinos', al agregar 'tigres' a la especificación, la extensión de esta nueva especificación sigue siendo una clase natural puesto que esta adición, en el contexto de nuestras teorías acerca del mundo, produce un peso explicativo adicional suficiente para explicar las propiedades y el comportamiento de los miembros de la extensión más restringida.

El nivel al que pertenece una palabra de clase natural se descubre *a posteriori*; depende de los resultados de nuestra investigación de la estructura básica de los objetos que forman la clase. Así, al investigar las formas explicativas de un conjunto de ejemplos paradigmáticos, podríamos descubrir dos o más formas explicativas distintas dentro de ese conjunto de ejemplificaciones. Se plantean entonces varias alternativas: a) Negar que se trate de una clase natural. b) Decir que sólo una de las formas explicativas define a la clase natural y excluir a los otros miembros del conjunto buscando para ellos una explicación y un nombre distinto. c) Decir que se trata de una clase natural de un nivel superior que incluye otras clases naturales de nivel inferior y que son éstas las que permiten explicar las estructuras básicas distintas. Sin embargo, para optar por esta tercera alternativa deberíamos descubrir rasgos comunes a las distintas formas explicativas que justificaran este paso.

Veamos por ejemplo el caso del jade que menciona Putnam. "Aun cuando los chinos no reconocen una diferencia, el término 'jade' se aplica a dos minerales: jadeita y nefrita. Químicamente hay una marcada diferencia. Jadeita es una combinación de sodio y aluminio. La nefrita está hecha de calcio, magnesio y hierro. Estas dos microestructuras tan distintas producen las mismas cualidades únicas de textura".⁹

¿Qué debemos hacer en este caso? ¿Decir que 'jade' no es una palabra de clase natural porque se refiere a dos tipos de objetos distintos? ¿O, seguir usando la palabra 'jade' para uno de los dos tipos y descartarla para el otro? ¿O, decidir por fin que se trata de una clase natural de nivel superior? Esta última alternativa no parece factible puesto que las dos formas explicativas que se han descubierto no parecen tener en común algo que pudiera explicar la similitud de la textura.

⁸ Platts, *op. cit.*

⁹ Putnam, *op. cit.*, p. 241

II

Una vez que hemos bosquejado esquemáticamente lo que son las clases naturales y cuáles son los términos con los que nos referimos a ellas, debemos preguntar si los términos mentales denotan o no clases naturales. Algunos filósofos contemporáneos han sostenido que la investigación neurofisiológica de las realizaciones¹⁰ de estados mentales nos permitirá descubrir su esencia o forma explicativa común. La tesis es que los estados mentales son clases naturales y que su esencia definitoria *a posteriori* puede formularse en términos de alguna ciencia natural, por ejemplo, la neurofisiología. Sin embargo, esta tesis deberá enfrentarse a objeciones muy serias que nos llevarán a concluir que, cualquiera que sea la esencia real de los estados mentales, no podrá ser física.

“Para que sea verdad que un estado mental ϕ (actitud proposicional o sensación) tenga una esencia física real ϕ , debe sostenerse que una criatura no puede ejemplificar ψ sin de esa manera también ejemplificar ϕ ; y *viceversa*; esto es, ‘ ψ ’ y ‘ ϕ ’ deberán ser necesariamente coextensivos.”¹¹ Pero esto no es así puesto que los estados mentales de un mismo tipo pueden realizarse físicamente de maneras muy distintas. La sensación de dolor, por ejemplo, sigue siendo dolor independientemente de la estructura neurofisiológica en que se dé. Dos criaturas distintas físicamente podrían realizar una sensación dolorosa mediante estados físicos diferentes.

En el caso de actitudes proposicionales parece todavía más obvio el que, por ejemplo, el deseo de tomar agua, o la creencia de que los ingleses son amables, puedan tener correlatos físicos muy distintos para personas diferentes o aún para la misma persona en ocasiones diversas.

Ahora bien, para nuestro propósito no es ni siquiera necesario que observemos casos reales de realización alternativa; bastaría con que existiera la posibilidad contrafáctica de realizaciones variables. Así, si alguien tiene un estado mental ψ , por ejemplo un dolor, y no está en estado cerebral ϕ , no diríamos por eso que no tiene dolor sino simplemente que realiza su dolor de una manera distinta.

Los correlatos físicos de los estados mentales no pueden fijar, mediante una definición *a posteriori*, la extensión de un término mental. Supongamos, por ejemplo, que el correlato físico del dolor fuese el estado cerebral B . Definimos al dolor como estado cerebral B , pero esta definición no puede fijar la extensión del término ‘dolor’ porque podría haber una sensación que caracterizamos como dolor, esto es, que es dolorosa y que, sin embargo, no fuera estado cerebral B . O, supongamos que un grupo de personas ejemplifican un estado mental —la creencia de que lloverá mañana— y que estipulamos que el

¹⁰ En el sentido de Putnam, *op. cit.*

¹¹ McGinn, *op. cit.*, p. 196.

estado mental que tiene la persona pertenece a la extensión del predicado si y sólo si su sistema nervioso satisface esa condición física que se correlaciona con el estado mental que ejemplifican esas personas. La consecuencia inaceptable sería que cualquier persona que realizara ese estado mental en forma distinta no podría ejemplificar ese estado mental. Cuando consideramos posibilidades contrafácticas, nos damos cuenta de que, aunque de hecho pudiera haber términos físicos y mentales coextensivos, no se trataría de una determinación definicional de la extensión.

Una posible manera de explicar las realizaciones distintas sería apelando a clases naturales de un nivel superior. Podemos decir, por ejemplo, que existen distintos *tipos* de dolor y que por eso hay realizaciones distintas. Esto es, que 'dolor' se refiere a una clase natural de un nivel superior y que, como tal, tiene una relación con los distintos dolores como la que tiene, por ejemplo, la clase de los metales con el oro y la plata. Pero, ¿cómo justificamos que el dolor sea una clase de nivel superior? En el caso de los metales *no* existen rasgos explicativos comunes que explican las características superficiales similares y distintas que tienen los diversos metales, pero ¿cuál es el contenido explicativo común de los distintos tipos de dolores? Por lo que hemos dicho antes, este elemento explicativo común no puede ser físico. Una referencia a la conducta tampoco parece adecuada si recordamos los problemas del reduccionismo conductista. No parece haber una manera adecuada de formular esta solución.

En general el que un término denote o no una clase natural se decide *a posteriori*; es epistémicamente posible que 'agua' no hubiese sido un término de clase natural. El que agua sea idéntico a H_2O también es algo que descubrimos *a posteriori*. En el caso de estados mentales, en cambio, y debido a la realización variable, sabemos *a priori* que los términos mentales no son términos de clases naturales determinables físicamente y que ningún tipo de sucesos mentales puede ser idéntico a algún tipo de sucesos físicos. Lo sabemos *a priori* debido a que los conceptos mentales obedecen a sus propios principios distintivos de aplicación; por ejemplo, las condiciones de inteligibilidad y racionalidad. La explicación psicológica está sujeta a la condición de que debemos interpretar al otro como un ser inteligible y racional. Esta condición conecta las atribuciones mentales con la conducta y no con estados físicos internos. No hay ninguna otra base para la atribución mental que no sea la conducta (verbal y no verbal) de una persona. Para poder decir que Juan está triste, debemos interpretar su conducta a la luz de sus creencias, deseos e intenciones.¹²

Si esto es así, entonces las realizaciones físicas de los estados mentales no pueden ser la base para que podamos atribuirle a alguien un estado mental y,

¹² Ver Davidson, *op. cit.*, ensayo 12.

por lo tanto, nuestro argumento muestra *a priori* que las leyes psicofísicas son imposibles y no sólo que son improbables o de hecho falsas.

Al comparar los términos de clases naturales con los términos de estados mentales nos damos cuenta de que funcionan de manera muy distinta en relación con otras características de las que habla Putnam.

Los términos de clases naturales exhiben un fenómeno que Putnam llama "la división de la tarea lingüística".¹³ Pertenecemos a una comunidad lingüística y hemos adquirido la palabra 'oro', pero al adquirir esta palabra no tenemos que adquirir también el método para reconocer si algo es o no oro. Para esto podemos recurrir siempre, en caso de duda, a un subgrupo de la comunidad, a los expertos, —en este caso a los químicos— que son los que conocen las definiciones que proporcionan condiciones necesarias y suficientes y los criterios de aplicación de la palabra.

"De esta manera el hecho más rebuscado acerca de (por ejemplo) agua, puede convertirse en parte del significado *social* de la palabra, a pesar de que lo desconozca la mayoría de los hablantes que adquieren la palabra."¹⁴

En el caso de predicados mentales no se da este fenómeno lingüístico. El hecho de que exista la posibilidad de realización variable no nos permite delegar de manera sistemática la atribución de estados mentales a especialistas en neurofisiología. Los términos mentales que aparecen en las explicaciones psicológicas del sentido común no podrán ser substituidos por descripciones científicas que descubrimos *a posteriori*. Nuestras descripciones psicológicas del sentido común definen a los estados mentales de una manera definitiva. Todos somos expertos en relación con la atribución de creencias, deseos, intenciones, sensaciones, etc. ❧

Nuestras clasificaciones de estados psicológicos no se afectan por el hecho de que descubramos realizaciones físicas distintas, ni tampoco parece que podamos descubrir una esencia física desconocida cuya revelación pudiera requerir una revisión de nuestro esquema psicológico ingenuo.

En el caso de términos de clases naturales en cambio, sí existe la posibilidad de que nuestras clasificaciones iniciales, basadas en apariencias superficiales, puedan reformarse, esto es, puede haber una reclasificación si así lo requiere nuestra mejor teoría científica. Al descubrir las formas explicativas básicas, podemos detectar casos que hemos clasificado erróneamente o ambigüedades. Tenemos por ejemplo el caso del oro falso o el de las dos sustancias distintas a las que nos referimos con la palabra 'jade'.

¹³ Putnam, *op. cit.*, pp. 227-229.

¹⁴ Putnam, *op. cit.*, p. 228. La división de la tarea lingüística no es esencial para un término de clase natural. Sí lo es, en cambio, el elemento 'deictico' (*indexical*).

III

Parece entonces que los términos mentales no se introducen en el lenguaje ni funcionan de la misma manera que los términos de clases naturales. Sin embargo, más allá de que estos términos funcionen o no como términos de clases naturales, nos interesa saber si efectivamente se refieren a objetos que forman clases naturales. Ya hemos visto que, debido a la realización variable, no podemos definir la esencia real de los estados mentales en términos de alguna ciencia física. Pero la pregunta que surge ahora es si los estados mentales tienen algún otro tipo de esencia real, esto es, si podemos identificar su esencia en términos que no sean físicos sino, por ejemplo, funcionales. ¿Hay clases naturales definidas funcionalmente? Un ejemplo podría ser el caso del páncreas.¹⁵ Parece ser que en este caso podemos descubrir su 'definición funcional' *a posteriori* mediante la investigación científica. Empezamos con casos paradigmáticos de páncreas y fijamos parcialmente la referencia de 'páncreas' en relación con estas ejemplificaciones. Después investigamos su constitución material y su funcionamiento y nos damos cuenta que el papel funcional común se realiza materialmente de distintas maneras. Sin embargo, dados nuestros intereses, decidimos que el papel funcional del páncreas tiene una fuerza explicativa suficientemente grande como para que podamos pasar por alto en este contexto las distintas realizaciones materiales. Por otro lado, la noción de un páncreas definido funcionalmente podría ocupar un lugar en una teoría explicativa más general acerca del funcionamiento de algunos organismos.

Hay otras clases de objetos que también definimos funcionalmente, por ejemplo, los lápices y las sillas. Éstos, sin embargo, no forman clases naturales puesto que en estos casos la definición no representa un descubrimiento empírico acerca de un tipo de objetos. No tenemos una teoría empírica acerca de las sillas; su definición es una especificación analítica, y aplicamos la palabra a los objetos que se adecuan a nuestra concepción de lo que debe ser y para lo que sirve una silla. En estos casos también encontramos el fenómeno de la realización variable, objetos físicamente distintos pueden tener la misma función, y es porque nos interesa ante todo esa función que los definimos en términos funcionales.

Sin embargo, la diferencia entre ambas definiciones es que en el caso de páncreas la definición en términos funcionales es *a posteriori* y tenemos una teoría explicativa que le confiere a la definición una fuerza explicativa suficiente que nos permite entender el que esos objetos tengan esas propiedades.

En el caso de silla, en cambio, no descubrimos qué es una silla mediante casos paradigmáticos, sino que estipulamos que llamaremos 'silla' al objeto que cumpla con tal y tal descripción.

¹⁵ Platts, *op. cit.*

¿Pero qué sucede con los estados mentales? ¿Forman éstos clases naturales cuya esencia real pueda definirse en términos funcionales? Veamos un ejemplo de definición funcional del dolor como “ese estado funcional *n* del sistema nervioso central que explica la tendencia del paciente, mientras persiste tal y tal estado en la conciencia, a evitar objetos que cree que son responsables de iniciar o agravar el estado”.¹⁶

Esta definición de la cual Wiggins dice que es “burda” o aproximativa, sin embargo, basta para nuestro propósito.

Supongamos entonces que la palabra ‘dolor’ se refiere a una clase natural; que fijamos su referencia mediante casos paradigmáticos de dolor y después descubrimos su forma explicativa. Aquí surge el primer problema, a saber, cómo identificar estos casos paradigmáticos de dolor en otras personas. ¿Los identificamos porque la persona *dice* que tiene un dolor? ¿Porque *actúa* de cierta manera? No parece que haya ninguna otra manera de identificar estados mentales en los otros si no es mediante la conducta (verbal y no verbal) de una persona.

Ahora bien, uno podría decir que la conducta nos permite identificar los casos de dolor de la misma manera como el estereotipo de agua (esto es, la descripción de los miembros típicos de la clase, en este caso, líquido, inodoro, insípido, transparente, etc.) nos permite identificar al principio los casos paradigmáticos de agua. Sin embargo, una vez que descubrimos la esencia real del agua —su estructura atómica— esta esencia real explicará los rasgos superficiales que se mencionan en el estereotipo. La idea es que hay aquí una relación necesaria entre esa forma explicativa común y cuando menos algunas de las características que aparecen en el estereotipo. Pero si esta es una relación de *explicación* debemos tener una manera de especificar la esencia real, que sea independiente de la manera en que especificamos el estereotipo.

En el caso del dolor, la conducta podrá hacer las veces del estereotipo que nos permite identificar casos paradigmáticos de dolor, y la descripción funcional sería la esencia real del dolor. Pero para que esto fuese así, el estado funcional debería *explicar* la conducta; y como la relación entre ellos sería necesaria,¹⁷ deberíamos poder especificar ese estado en términos independientes de la conducta. Sin embargo, la definición de Wiggins no cumple con esa condición. No podemos caracterizar al estado funcional *n* de una manera independiente, esto es, no descubrimos nada más acerca de *n* aparte de que es el

¹⁶ David Wiggins, “Identity, Designation, Essentialism and Physicalism” en *Philosophia*, Vol. 5. enero-abril 1975, p. 23.

¹⁷ Si la relación entre conducta y estado funcional fuese contingente, entonces no se trataría de una relación de *explicación*, que es lo que debería ser si se trata de una clase natural. Por otro lado, si decimos que alguien tiene un dolor pero no tiene ninguna tendencia a algún tipo de conducta, no habría manera de identificarlo.

responsable de que tengamos la tendencia a cierta conducta: a evitar objetos que inician o agravan ese estado de dolor.

Atribuimos dolor, esto es, atribuimos *n*, con base en una conducta y después pretendemos que *n* es el que explica la tendencia a esa conducta.

Para que podamos tener una verdadera explicación en términos funcionales, es necesario que el *explanandum* y el *explanans* puedan formularse de manera independiente. Mientras no podamos dar una definición funcional de los estados mentales sin hacer mención a la conducta, no podemos decir que se trata de una especificación de la esencia real de una clase natural.

Aun si el funcionalista logra especificar el estado mental de una manera independiente de la conducta, podría surgir aquí otro problema en relación con las definiciones funcionales: el problema del holismo. La conexión entre estado mental y conducta opera de una manera holista. Para atribuirle a alguien un estado mental debo interpretar su conducta a la luz de una serie de otros conjuntos de estados mentales (en nuestro caso, la persona *Cree* que ese objeto es responsable de su dolor, *reconoce* al objeto como tal, etc.) No pueden describirse las circunstancias en las que se manifiesta la tendencia a cierta conducta sin introducir otros conceptos mentales. Esto quizá no sea un problema para el funcionalista. Lo que sería un problema es que el contenido de la disposición —aun especificado de la manera más general— dependiera de otros estados mentales.

Estos comentarios esquemáticos acerca de las relaciones entre estados mentales y estados funcionales nos dan una idea de por qué los estados mentales *no* forman clases naturales cuya esencia real pueda especificarse en términos funcionales. Si no podemos especificar una esencia real para estados mentales en términos físicos ni en términos funcionales (ni de ninguna otra manera), debemos concluir que los estados mentales no tienen ningún tipo de esencia real y, por lo tanto, que *no* son clases naturales.

IV

En esta parte me ocuparé de algunas de las razones para sostener la primera premisa del argumento, a saber, que los términos que aparecen en una ley deberán ser términos de clases naturales.

A diferencia de las generalizaciones verdaderas pero accidentales como “todas las monedas que están en mi bolsillo son de plata”, las generalizaciones que sí pueden considerarse como leyes deberán cumplir ciertas condiciones. Así, una ley deberá apoyar condicionales contrafácticos y subjuntivos; deberá ser confirmable por sus instancias positivas y servir para hacer predicciones.¹⁸ El que una generalización apoye o no condicionales contrafácticos depende

¹⁸ Son contrafácticos los condicionales en los que se supone que el antecedente no es el caso, por ejemplo, “si *x* hubiera sido *A*, entonces habría sido *B*”. Son subjuntivos los condicionales que

a su vez de que sea proyectible a casos no observados, futuros y no actuales. La proyectibilidad de una ley está en función de los predicados que aparecen en ella; esto es, hay predicados que son proyectibles y otros que no lo son.¹⁹ Una manera de solucionar el problema de distinguir a los predicados que *sí* son proyectibles es diciendo que se trata de predicados de clases naturales. Así, una de las condiciones que debe cumplir una generalización para ser una ley, es que los predicados que aparezcan en ella sean palabras de clases naturales, es decir, que por lo menos sean verdaderos de todos y sólo los objetos que forman una clase natural.

Una de las maneras de mostrar la importancia de esta tesis es que sirve para solucionar paradojas como la de las esmeraldas verdes de Goodman²⁰ y la paradoja de los cuervos de Hempel. Veamos esta segunda paradoja:²¹

Todo cuervo negro sirve para confirmar la ley que dice que todos los cuervos son negros. Pero, si esto es así, entonces toda hoja verde que es no-negra y no-cuervo debería servir para confirmar la ley de que todos los objetos no-negros son no-cuervos, esto es, que todos los cuervos son negros. La paradoja consiste en que una hoja verde pueda servir para confirmar la ley de que todos los cuervos son negros.

Ahora bien, 'cuervo' y 'negro' son predicados proyectibles porque son predicados de clases naturales. Pero el complemento de un predicado proyectible no tiene por qué ser también proyectible. En efecto, 'no-negro' y 'no-cuervo' no son proyectibles puesto que los objetos que no son negros y los que no son cuervos no forman clases naturales. Esto significa entonces que "todos los cuervos son negros" es una ley y "todos los objetos no-negros son no-cuervos" —aunque lógicamente equivalente a la anterior— no lo es puesto que sus predicados no son términos de clases naturales. La equivalencia lógica no preserva la legalidad. Ahora bien, aunque la confirmación es independiente de la legalidad, esto es, que hay enunciados no-legales que son confirmados por sus instancias positivas, esto no quiere decir que las mismas instancias

afirman que "si x fuera A , entonces sería B ", pero se deja abierta la posibilidad de que x , de hecho sea o no sea A . Un enunciado de la forma "Todas las A son B ", sólo puede ser considerado una ley si apoya estos tipos de condicionales y si sirve como base para una explicación o predicción. Por ejemplo, "Todo trozo de sodio se quema si se le expone al aire" es una ley puesto que puede explicar por qué un trozo de sodio se quema si se le expone al aire, la utilizamos también para predecir que un trozo de sodio se quemará si se le expone al aire y además, apoya el condicional contrafáctico que afirma que un trozo de sodio que no ha sido expuesto al aire se quemaría si se le expusiera.

¹⁹ Ver Goodman, *Fact, Fiction and Forecast*, Bobb & Merrill, 1975 p. 94.

²⁰ Goodman, *op. cit.*, pp. 87–108 y mi "¿Existen las leyes psicofísicas?", en *Teoría*, FFL, UNAM, 1980.

²¹ Hempel, *Aspects of Scientific Explanation and Other Essays*, Free Press, Nueva York, 1965, p. 15. Ver también W. O. Quine, "Natural Kinds" en *Theories and Things*, Harvard University Press. 1981. Traducción al español: *Teorías y cosas*, UNAM, México, 1986.

confirman enunciados lógicamente equivalentes. Así, los cuervos negros sirven para confirmar el enunciado “todos los cuervos son negros” pero no para confirmar “todos los objetos no-negros son no-cuevos”. Instancias confirmatorias de este último serían los objetos no-negros y no-cuevos, sin embargo, dada la amplitud de esta clase, la confirmación que estos objetos proporcionarían sería muy débil. La equivalencia lógica no preserva ni la legalidad ni las relaciones de confirmación.

En el caso de leyes como “Para toda x , si x es cobre entonces x es conductor de electricidad”, los predicados ‘ser cobre’ y ‘conductor de electricidad’ denotan clases naturales. ‘Cobre’ es una palabra de clase natural y las cosas de cobre forman una clase natural puesto que tienen una esencia real o forma explicativa común. Por otro lado, los objetos que son conductores de electricidad comparten también una forma explicativa básica, y constituyen, por tanto, una clase natural aunque de un nivel superior.

Aquí cabría mencionar otra tesis acerca de las leyes naturales: “para que una generalización pueda considerarse una ley, las propiedades especificadas por los predicados que contiene no sólo deberán estar universalmente coinstanciadas, sino que deberán estar conectadas de cierta manera”.²²

Una explicación plausible de esta conexión es que la coinstanciación observada sea el resultado de la presencia de algún mecanismo o estructura generativa subyacente cuya elucidación detallada serviría para explicar la coincidencia de propiedades.²³

Así, “Todo cobre es conductor de electricidad” es una ley porque el cobre es una clase natural cuya esencia real o forma explicativa básica es la responsable de que el cobre sea conductor de electricidad y, en la medida en que tengamos una teoría que nos permita precisar esta esencia real, podremos dar una explicación teórica de por qué el cobre es conductor de electricidad.

Esto no es así en el caso de las generalizaciones psicofísicas puesto que los predicados mentales no denotan clases naturales cuya esencia real pudiera ser la responsable de, y sirviera para, explicar la co-instanciación de una propiedad mental y una propiedad física. Pero, si los términos mentales no denotan clases naturales, las generalizaciones en que aparecen predicados mentales no podrán ser leyes.

Pero veamos tipos distintos de leyes psicofísicas. En primer lugar están las leyes-puente cuyo propósito es la reducción de predicados mentales a predicados físicos. Podemos descartar desde un principio a los enunciados que expresan identidades de propiedades mentales y físicas puesto que ya hemos visto que estas identidades no existen. Pero podemos —si dejamos de

²² McGinn, *op. cit.*, p. 205.

²³ Esto, sin embargo, no funcionaría en el caso de que tuviéramos leyes básicas que se refirieran a partículas no estructuradas, porque si no hay estructura no puede haber tampoco un mecanismo generativo.

lado presupuestos ontológicos— considerar bicondicionales legales verdaderos que contienen un predicado mental ' ψ ' y un predicado físico ' ϕ ', esto es, $(x) (\psi x \leftrightarrow \phi x)$ o, por ejemplo, dolor si y sólo si excitación de Fibras-C.

Ahora bien, un enunciado legal deberá ser proyectible. Sin embargo, y debido a la realización variable, la observación de un subconjunto de casos a los que se aplica la ley, *no* nos permite proyectarla a casos no observados o no actuales puesto que existe la posibilidad de que haya criaturas que realicen sus estados mentales de maneras distintas. Esta generalización podría ser verdadera de hecho, podría conformarse a todos los casos actuales, pero no podría apoyar condicionales contrafácticos, esto es, no podríamos decir que si algo fuese dolor entonces sería excitación de Fibras-C, ni tampoco que si algo fuese excitación de Fibras-C entonces sería dolor.

Una ley psicofísica que sea un bicondicional requiere que fijemos tanto la extensión de los términos mentales mediante términos físicos, como la extensión de los términos físicos mediante términos mentales. Es decir $(x) (\psi x \leftrightarrow \phi x)$ implica que sean verdaderos tanto $(x) (\psi x \rightarrow \phi x)$ como $(x) (\phi x \rightarrow \psi x)$.

Ahora bien, debido a la realización variable, los correlatos físicos de los estados mentales no pueden fijar, mediante una definición *a posteriori*, la extensión de un término mental. El que alguien tenga un estado mental dado no determina necesariamente cuál sea el estado físico que lo realice.

El caso converso, con un antecedente físico y un consecuente mental, parece más difícil de rechazar debido a la tesis de la superveniencia de lo mental a lo físico que afirma que no puede haber dos sucesos o estados iguales en todos sus aspectos físicos y diferir en un aspecto mental. Así, el que una criatura esté en un estado físico ϕ que realiza un estado mental ψ , parece implicar que cualquier otra criatura que esté en ϕ estará también en ψ . Supongamos, por ejemplo, que el estado cerebral B es la realización física del dolor que tiene Juan. Si examinamos después a Pedro y éste se encuentra también en el estado cerebral B , esto nos permitirá concluir, aparentemente, que él también tiene dolor.

Sin embargo, es necesario aclarar que la tesis de la superveniencia es una tesis general, esto es, no podemos tomar en cuenta a ϕ solo, sino que hay que asumir, además, que no hay otra diferencia pertinente en los estados del cerebro. Así, Pedro puede estar en el estado cerebral B pero, debido a alguna otra diferencia física relevante, podría no sentir dolor. Pero si esto es así, entonces no podemos decir ya que para todos los casos, si alguien está en estado ϕ entonces estará en estado ψ .

Un condicional de superveniencia, para que fuera confirmado por sus instancias positivas y apoyara proyecciones subjuntivas, debería formularse de tal manera que la condición física incluyera todos los aspectos relevantes del cerebro en los cuales superviene ψ . Ahora bien, dada la realización variable,

es posible que haya varios condicionales de superveniencia que tengan como consecuente una condición mental dada, y esto crea algunas dificultades generales en relación con este tipo de condicionales:

- a) "... conocer la verdad de un número finito de estos condicionales contiene un consecuente mental, no proporciona una manera efectiva de identificar el conjunto de predicados físicos en el cual podría en principio supervenir el predicado mental."²⁴ Esto es, debido a que un predicado mental puede en principio realizarse físicamente de muchas maneras, el conocer una o algunas de estas posibles realizaciones, no nos permite identificar el conjunto de realizaciones del predicado mental. Siempre queda abierta la posibilidad de que un predicado mental tenga una realización física distinta. Un condicional de este tipo sólo nos dice que si alguien realiza su dolor mediante un estado físico ϕ , y asumiendo que no hay otra diferencia pertinente en relación con otros estados cerebrales, entonces todos los otros casos en los que se dé ϕ , serán también casos de dolor. Pero esto no implica que ϕ sea la única realización posible del dolor.
- b) De la posibilidad de que haya distintas condiciones físicas que realicen una condición mental dada se sigue que "... no puede darse una explicación general, en términos físicos, de la condición mental y de los sucesos mentales que esa condición circunscribe."²⁵ No puede darse una explicación general, porque distintas realizaciones físicas pueden no tener nada en común que permita su unificación teórica y, porque siempre existe la posibilidad de que surja una realización física que no habíamos tomado en cuenta.

Aquí me parece pertinente el siguiente comentario: lo que está en discusión es si las generalizaciones psicofísicas son o no leyes estrictas. No estoy poniendo en duda que existan de hecho algunas generalizaciones psicofísicas que sean sumamente confiables. Es posible, por ejemplo, que para nosotros y dado como somos, existiera una manera natural de realizar el dolor. "Natural" querría decir aquí, dada la evolución del ser humano, dada la manera en que de hecho se ha desarrollado biológicamente la especie. Así decimos, por ejemplo, que la capacidad que tiene una criatura de que su sangre circule en su sistema circulatorio, se realiza naturalmente por el hecho de que la criatura tenga un corazón.²⁶

Sin embargo, las generalizaciones psicofísicas en las que se especifican realizaciones físicas naturales, parecen más plausibles para el caso de estados mentales que tienen características fenomenológicas como las sensaciones o

²⁴ McGinn, *op. cit.*, pp. 207-08.

²⁵ *Ibid.*, p. 208.

²⁶ J. Hopkins, "Mental States, Natural Kinds and Psychophysical Laws" II, *op. cit.*, p. 227.

los casos de percepción sensorial. En el caso de actitudes proposicionales, en cambio, no parece haber realizaciones naturales, puesto que las creencias, deseos, etc., pueden formarse de múltiples maneras y cumplir papeles muy distintos en su relación con la conducta. En estos casos suponemos más bien que las realizaciones podrían ser distintas para cada persona y hasta para una misma persona en distintas circunstancias.

Pero aun cuando fuese en principio posible establecer una condición física suficiente para atribuirle a alguien, por ejemplo, el predicado “cree que Beethoven murió en Viena”, esto no nos proporciona un sustituto de fuerza explicativa equivalente a las condiciones que podemos expresar en términos psicológicos, es decir, tomando en cuenta otras creencias, intenciones, observando sus acciones, etc., de tal manera que la atribución de esa nueva creencia esté de acuerdo con una interpretación que hacemos de él como persona.

Examinemos, en segundo lugar, las leyes psicofísico causales.

Una primera objeción general que podemos hacer a este tipo de leyes es que, debido a la realización variable de los sucesos mentales, no podemos aceptar que exista una ley que afirme que un mismo tipo de sucesos mentales esté causado por un mismo tipo de sucesos físicos dado que las causas físicas varían de acuerdo con las distintas realizaciones físicas que tenga un mismo tipo de sucesos mentales.

Debido a la posibilidad de una realización variable, es perfectamente posible que nuestra manera de clasificar a los sucesos cuando los describimos en términos físicos y de tal manera que puedan ejemplificar una ley física, no corresponda a la manera en la que los clasificamos cuando los describimos en términos psicológicos.

Una posible formulación de una ley causal es la que propone Davidson:²⁷

$$(S) (e)(n) ((Fe \ \& \ t(e)=n \rightarrow (\exists !f) (Gf \ \& \ t(f)=n + \in \ \& \ C(e,f)))$$

(L) y

$$(N) (e)(n) ((Ge \ \& \ t(e)=n + \in) \rightarrow (\exists !f) (Ff \ \& \ t(f)=n \ \& \ C(f,e)))$$

en donde las variables ‘e’ y ‘f’ cuantifican sobre sucesos, ‘n’ sobre números; F y G son propiedades de sucesos; C (e,f) debe leerse como ‘e causa f’ y ‘t’ es una función que asigna un número a un suceso para señalar el tiempo en el que ocurre el suceso.

Según esta formulación las leyes causales consisten en una conjunción de condicionales que cuantifican sobre sucesos particulares; uno de los condicionales afirma suficiencia, el otro necesidad.

Una pregunta que debemos hacernos es si podemos formular —apoyándonos en la tesis de la superveniencia— una condición causal suficiente en

²⁷ Davidson, “Casual Relations” en *op. cit.*, ensayo 7.

términos físicos para un tipo dado de sucesos mentales. Esto es, si suponemos que hay una ley que afirma que los sucesos físicos del tipo ϕ son causados por sucesos físicos del tipo ϕ' . Y, si hay superveniencia de ψ en ϕ con respecto a un suceso particular s , entonces se sigue causalmente que si ocurre un suceso del tipo ϕ' , ocurrirá también un suceso del tipo ψ .

El problema con esta condición causal suficiente es que está sujeta a las mismas dificultades que mencioné en relación con los condicionales de superveniencia.

Pero existe también un problema en relación con la condición de necesidad. De nuevo, debido a que una condición mental dada puede supervenir en varias condiciones físicas, no es posible formular una condición física necesaria. Los sucesos mentales circunscritos por una propiedad mental dada pueden ser causados de maneras muy distintas, y es posible que no podamos encontrar una descripción física que sea común a las distintas causas, y que sea por tanto, la que nos proporcione la condición de necesidad.

Si la discusión anterior es correcta, debemos concluir que las descripciones psicológicas no pueden aparecer en leyes estrictas o, lo que es lo mismo, que no existen las leyes psicofísicas.