

Block, Ned (1996) "What is functionalism?" *The Encyclopedia of Philosophy Supplement*.

El funcionalismo

Ned Block

¿Qué es el funcionalismo? El funcionalismo es una de las propuestas más importantes que se ha ofrecido como solución del problema mente-cuerpo. Las soluciones del problema mente-cuerpo normalmente tratan de responder cuestiones como: ¿Cuál es la naturaleza última de lo mental? En el nivel más general, ¿qué es lo que hace mental a un estado mental? O más específicamente, ¿qué tienen en común los pensamientos en virtud de lo cual son pensamientos? Esto es, ¿qué es lo que hace a un pensamiento un pensamiento? ¿Qué es lo que hace a un dolor un dolor? El dualismo cartesiano ha dicho que la naturaleza última de lo mental se hallaba en una sustancia mental especial. El conductismo identificó los estados mentales con disposiciones conductuales; el fisicismo en su versión más influyente identifica los estados mentales con estados cerebrales¹. El funcionalismo dice que los estados mentales están constituidos por sus relaciones causales entre sí y con los *inputs* sensoriales y *outputs* conductuales². El funcionalismo es uno de los desarrollos teóricos más importantes de la filosofía analítica del siglo veinte y proporciona el apuntalamiento conceptual de muchos trabajos en ciencia cognitiva.

El funcionalismo tiene tres fuentes distintas. Primero, Putnam y Fodor vieron los estados mentales en términos de una teoría computacional empírica de la mente. Segundo, los análisis "tópicamente neutrales" [*topic neutral*] de Smart condujeron a Armstrong y Lewis a un análisis funcionalista de los conceptos mentales. Tercero, la idea de Wittgenstein del significado como uso condujo a una versión de funcionalismo como teoría del significado, además desarrollada por Sellars y posteriormente Harman.

Una motivación por detrás del funcionalismo puede comprenderse atendiendo a conceptos de aparatos como *carburador* y conceptos biológicos como *riñón*. Lo que hace de algo un carburador es mezclar combustible con aire en un motor de combustión interna –*carburador* es un concepto funcional–. En el caso del riñón, el concepto *científico* es funcional –definido en términos de un rol en la filtración de la sangre y la conservación de ciertos balances químicos–.

¹ N.T. De acuerdo con la decisión terminológica adoptada por unanimidad en el II Congreso de la Sociedad Española de Filosofía Analítica llevado cabo en la Universidad Complutense de Madrid en 1998, traduzco el término inglés 'physicalism' por 'fiscicismo' y no por 'fiscalismo'. El mismo criterio adopto para el término inglés 'externalism', el cual traduzco por 'externismo' y no por 'externalismo'. Cf. Lowe, Edward J., *Filosofía de la mente*, Idea Books, Barcelona, 2000, Nota de la traductora.

² N.T. Se trata de *entradas sensoriales* y *salidas conductuales*, utilizo la terminología de *inputs* y *outputs* en concordancia con la traducción de Eleonora Baringoltz en Block, Ned, "Las dificultades del funcionalismo" (selección), en Rabossi, Eduardo (compilador), *Filosofía de la mente y ciencia cognitiva*, Paidós, Barcelona, 1995.

El tipo de función relevante para la mente puede ser introducido por vía del autómata detector de paridad ilustrado en la figura de abajo, el cual nos dice si ha visto un número par o impar de '1s'. Este autómata tiene dos estados, S_1 y S_2 ; un *input*, '1' (aunque su *input* pueda ser nada) y dos *outputs*, pronuncia la palabra "par" o "impar". La tabla describe dos funciones, una desde el *input* y el estado hasta el *output*, y otra desde el *input* y el estado hasta el próximo estado. Cada casilla codifica dos condicionales especificando el *output* y el próximo estado, dados el estado actual y el *input*. La casilla de la izquierda dice que si la máquina está en S_1 y ve un '1', dice "impar" (indicando que ha visto un número impar de '1') y va hasta S_2 . La casilla de la derecha dice, de manera similar, que si la máquina está en S_2 y ve un '1', dice "par" y vuelve a S_1 .

S_1	S_2
"Impar"	"Par"
S_2	S_1

Ahora supongamos que preguntamos: "¿Qué es S_1 ?". La respuesta es que la naturaleza de S_1 es completamente relacional y está completamente capturada por medio de la tabla. Podríamos dar una caracterización explícita de S_1 de la siguiente manera:

Estar en S_1 = Estar en el primero de dos estados que están relacionados entre sí y con los *inputs* y *outputs* de la siguiente manera: estar en uno de los estados y obteniendo un *input* de '1' ocasiona entrar en el segundo estado y proferir "impar"; y estar en el segundo de los dos estados y obtener un *input* de '1' ocasiona entrar en el primer estado y proferir "par".

Haciendo la cuantificación sobre estados más explícita:

Estar en S_1 = Ser un x tal que $\exists P \exists Q$ [si x está en P y obtiene un *input* de '1', entonces entra en Q y profiere "impar"; si x está en Q y obtiene un *input* '1', entra en P y profiere "par" & x está en P] (Nota: léase ' $\exists P$ ' como hay una propiedad P).

Esta ilustración puede ser utilizada para puntualizar algunas cosas. (1) De acuerdo con el funcionalismo, la naturaleza de un estado mental es como la naturaleza de un estado de autómata: constituido por sus relaciones con otros estados y con los *inputs* y *outputs*. S_1 no es más que estar [en un estado] tal que recibiendo un *input* de '1' ocasiona tal y tal cosa, etc. De acuerdo con el funcionalismo, tener dolor no es más que disponerte a decir 'ouch', preguntarte si estás herido, distraer tu atención, etc. (2) Puesto que los estados mentales son como estados de autómata en este sentido, el método ilustrado para definir estados de autómata se supone que también funciona para estados mentales. Los estados mentales

pueden ser totalmente caracterizados en términos que involucren únicamente lenguaje lógico-matemático y términos para señales de *input* y *outputs* conductuales. Por lo tanto el funcionalismo satisface una de las desiderata del conductismo, caracterizar lo mental en un lenguaje completamente no mental. (3) S_1 es un estado de segundo orden que consiste en tener *otras* propiedades, digamos, propiedades mecánicas, hidráulicas o electrónicas que tienen ciertas relaciones entre sí. Estas otras propiedades, cuantificadas sobre las definiciones recién dadas, se dice que son *realizaciones* de las propiedades funcionales. Por lo tanto, aunque el funcionalismo caracteriza lo mental en términos no mentales, lo hace sólo por medio de la cuantificación sobre realizaciones de estados mentales, lo cual no habría alegrado a los conductistas. (4) Un estado funcional puede ser realizado de diferentes maneras. Por ejemplo, una misma máquina de metal y plástico que satisfaga la tabla de la máquina puede ser hecha de engranajes, ruedas, poleas y cosas por el estilo, en tal caso la realización de S_1 sería un estado mecánico; o la realización de S_1 podría ser un estado electrónico, y así sucesivamente. (5) De la misma manera que un estado funcional puede ser realizado de diferentes maneras, un estado físico puede realizar diferentes estados funcionales en diferentes máquinas. Esto podría suceder, por ejemplo, si un mismo tipo de transistor fuese utilizado para hacer diferentes cosas en diferentes máquinas. (6) Mientras S_1 puede ser realizado de muchas maneras, la afirmación de que S_1 es un estado mecánico sería falsa (podría decirse al menos), como lo sería la afirmación de que S_1 es un estado electrónico. Por esta razón, hay un caso importante en que el funcionalismo muestra que el fisicismo es falso: si una criatura sin cerebro puede pensar, el pensamiento no puede ser un estado cerebral. (Pero véase la sección acerca de funcionalismo y fisicismo más abajo).

La noción de realización merece una discusión adicional. En los inicios del funcionalismo, una propiedad de primer orden a menudo se consideraba que realizaba una propiedad funcional en virtud de una correspondencia 1-1 entre los dos reinos de propiedades. Pero una definición tal de realización produce demasiadas realizaciones. Supóngase, por ejemplo, que en t_1 gritamos ‘uno’ a un balde de agua, y entonces en t_2 gritamos ‘uno’ nuevamente. Podemos considerar al balde como un autómatas detector de paridad haciendo corresponder la configuración física del balde en t_1 con S_1 y el calor emitido o absorbido por el balde en t_1 con “impar”; haciendo corresponder la configuración física del balde en t_2 con S_2 y el calor intercambiado con el ambiente en t_2 con “par” y así sucesivamente. Lo que se deja fuera en este modo de correlación *post hoc* del pensamiento de realizaciones es que una realización verdadera debe satisfacer los *contrafácticos* implícitos en la tabla. Para ser una realización de S_1 , no es suficiente con que conduzca a un cierto *output* y estado, dado un *input* de ‘1’; también se requiere que teniendo el *input* ‘0’, la realización S_1 conduzca a otro *output* y estado. La satisfacción de los contrafácticos relevantes es construida dentro de la noción de realización mencionada arriba en (3). Véase Lycan (1987).

Supóngase que tenemos una teoría de los estados mentales que especifica todas las relaciones causales entre estados, *inputs* sensoriales y *outputs* conductuales. Enfocando el dolor como un ejemplo de estado mental, podría decirse, entre otras cosas, que sentarse sobre una tachuela causa dolor y que el dolor causa preocupación y decir ‘ouch’. Aceptando, a los efectos del ejemplo, proseguir con esta teoría tonta, el funcionalismo diría entonces que podemos definir

‘dolor’ de la siguiente manera: tener dolor = estar en el primero de dos estados, el cual es causado por sentarse sobre tachuelas y a su vez causa el otro estado y la preferencia de ‘ouch’. De una manera más simbólica:

Tener dolor = Ser un x tal que $\exists P \exists Q$ [sentarse sobre una tachuela causa P & P causa tanto Q como la preferencia de ‘ouch’ & x está en P]

De una manera más general, si T es una teoría psicológica con n términos mentales de los cuales el 17 es ‘dolor’, podemos definir ‘dolor’ en relación con T de la siguiente manera (‘F₁’... ‘F_n’ son variables que reemplazan los términos mentales n, e₁, etc. y o₁, etc. indica):

Tener dolor = Ser un x tal que $\exists F_1 \dots \exists F_n$ [T (F₁... F_n, i₁, etc., o₁, etc.) & x es en F₁₇]

De esta manera, el funcionalismo caracteriza lo mental en términos no mentales, en términos que involucran cuantificación sobre realizaciones de estados mentales pero sin hacer mención explícita de ellos; por lo tanto el funcionalismo caracteriza lo mental en términos de estructuras que están fijadas a la realidad únicamente a través de los *inputs* y *outputs*.

La teoría psicológica T recién mencionada puede ser una teoría psicológica empírica o bien una teoría “popular” de sentido común y los funcionalismos resultantes son muy diferentes. En el último caso, el *funcionalismo conceptual*, las definiciones funcionales tienen por objetivo capturar nuestros conceptos mentales corrientes. En el primer caso, que denominé *psicofuncionalismo*, las definiciones funcionales no se supone que capturen los conceptos corrientes sino únicamente se supone que fijen las extensiones de los términos mentales. La idea del psicofuncionalismo es que la naturaleza científica de lo mental no consiste en algo biológico sino en algo “organizacional”, análogo a la estructura computacional. El funcionalismo conceptual, por el contrario, puede ser pensado como un desarrollo del conductismo lógico. Los conductistas lógicos pensaban que el dolor era una disposición a tener *conducta* de dolor. Pero como han señalado Geach y Chisholm, lo que cuenta como conducta de dolor depende de las creencias y deseos del agente. Los funcionalistas conceptuales evitan este problema definiendo cada estado mental en términos de su contribución a las disposiciones de comportarse y tener otros estados mentales.

Funcionalismo y fisicismo. Las teorías de la mente previas al funcionalismo han estado ocupadas de (1) qué *es* y (2) qué es lo que da a cada tipo de estado mental su identidad propia, por ejemplo, qué tienen en común los dolores en virtud de lo cual son dolores. Desplegando estos términos un poco, podemos decir que (1) es una cuestión ontológica [*matter of ontology*] y (2) [es una cuestión] metafísica [*of metaphysics*]. He aquí las afirmaciones ontológicas: el dualismo nos dijo que hay tanto sustancias mentales como físicas, mientras que el conductismo y el fisicismo son monistas, afirmando que hay únicamente sustancias físicas. He aquí las afirmaciones metafísicas: el conductismo nos dice que lo que los dolores (por ejemplo) tienen en común en virtud de lo cual son dolores es algo conductual; el dualismo dio una respuesta no-física a esta cuestión y el fisicismo da una respuesta física a esta cuestión. Volviendo ahora al funcionalismo, éste responde la cuestión metafísica [*metaphysical question*] sin responder la cuestión ontológica [*ontological question*]. El

funcionalismo nos dice que lo que los dolores tienen en común –lo que los hace dolores– es su función; pero el funcionalismo no nos dice si los seres que tienen dolores tienen alguna parte no-física. Este punto puede ser visto en términos del autómata descrito más arriba. Para ser una autómata del tipo descrito, una máquina concreta real necesita solamente tener estados relacionados entre sí y con los *inputs* y los *outputs* de la manera descrita. La descripción de la máquina no nos dice cómo funciona la máquina o de qué está hecha, y en particular no descarta una máquina que esté operada por un alma inmaterial, con tal de que el alma pueda operar de la manera determinista especificada en la tabla. Véase Putnam, *op. cit.* y el artículo de Fodor en Block (1980).

Pensando en la relación entre el funcionalismo y el fisicismo, es útil diferenciar dos categorías de tesis fisicistas (Véase FISCISMO): Una versión de fisicismo compite con el funcionalismo, haciendo una afirmación metafísica acerca de la naturaleza física de las propiedades de los estados mentales o tipos (y es por consiguiente a menudo llamada fisicismo de “tipos” [*“type” physicalism*]). Como se ha mencionado más arriba, desde un punto de vista, el funcionalismo muestra que el fisicismo de tipos es falso.

Sin embargo, hay fisicismos más modestos cuyo impulso es ontológico más que metafísico. Tales afirmaciones fisicistas no son en absoluto incompatibles con el funcionalismo. Considere, por ejemplo, un fisicismo que dice que toda cosa real está hecha enteramente de partículas del tipo que componen la materia inorgánica. En este sentido de fisicismo, la mayoría de los funcionalistas han sido fisicistas. Además, el funcionalismo puede ser modificado en una dirección fisicista, por ejemplo, requiriendo que todas las propiedades cuantificadas en una definición funcional sean propiedades físicas. El fisicismo de tipos es a menudo contrastado con el fisicismo de *casos* [*token physicalism*]. (La palabra ‘diente’ en esta oración tiene seis casos [*token*] de letras pero cinco tipos [*types*] de letras). El fisicismo de casos dice que cada dolor (por ejemplo) es un estado físico, pero el fisicismo de casos permite que pueda no haber algo físico que todos los dolores compartan, nada físico que haga de un dolor un dolor.

Es una peculiaridad de la literatura acerca del funcionalismo y del fisicismo que mientras algunos funcionalistas digan que el funcionalismo muestra que el fisicismo es falso (Véase los artículos de Putnam, Fodor, y Block y Fodor en Block (1980), algunos de los cuales están también en otras antologías), otros digan que el funcionalismo muestra que el fisicismo es verdadero (Véase los artículos de Lewis y Armstrong en Block (1980) y Rosenthal (1991)). En el caso de Lewis, el asunto es en parte terminológico. Lewis es un funcionalista conceptual acerca de *tener dolor*. ‘Tener dolor’ en la reglamentación de Lewis, podría decirse que es un designador rígido de una propiedad funcional. [Un designador rígido nombra la misma cosa en cada mundo posible. ‘El color del cielo’ es no-rígido, ya que nombra al rojo en mundos en los cuales el cielo es rojo. ‘Azul’ es rígido, ya que nombra al azul incluso en mundos donde el cielo es rojo]. ‘Dolor’, por contraste, es un designador no-rígido conceptualmente equivalente a una descripción determinada de la forma ‘el estado con tal y tal rol causal’. El referente de esta frase en nosotros, sostiene Lewis, es un cierto estado cerebral, aunque el referente de esta frase en un robot podría ser un estado de un circuito, y el referente en un ángel sería un estado no-físico. De manera similar, ‘el número ganador’ remite a ‘17’ en una lotería y a ‘596’ en

otra. Por tanto, Lewis es un funcionalista (un funcionalista conceptual por cierto) acerca de tener dolores. En términos de la cuestión metafísica descrita más arriba –qué tienen en común los dolores en virtud de lo cual son dolores–, Lewis es un funcionalista, no un fisicista. Lo que comparten mis dolores y los dolores del robot es un rol causal, no algo físico. Así como no hay similitud numérica entre 17 y 596 más que ser números ganadores, no hay similitud física de ningún tipo entre los dolores humanos y los dolores de los ángeles. Sin embargo, acerca de la cuestión de la naturaleza científica del dolor, Lewis es un fisicista. Lo que tienen en común el dolor humano y marciano desde su perspectiva es algo conceptual, no algo científico.

El funcionalismo y las actitudes proposicionales. La discusión de la caracterización funcional proporcionada más arriba supone una teoría psicológica con un número finito de términos de estados mentales. En el caso de estados monádicos como el dolor, la sensación de rojo, etc., parece ser una opción teórica simple listar los estados y sus relaciones con otros estados, *inputs* y *outputs*. Pero por varias razones, esto no es una opción teórica sensata para estados-de-creencia, estados-de-deseo y otros estados de actitudes proposicionales. Para empezar, la lista sería demasiado larga para ser representada prescindiendo de métodos combinatorios. Ciertamente podría decirse que no hay límite superior en cuanto al número de proposiciones que podrían ser en principio un objeto de pensamiento. Por otro lado, hay relaciones sistemáticas entre las creencias: por ejemplo, la creencia de que Juan ama a María y la creencia de que María ama a Juan. Estos estados de creencias representan los mismos objetos pero relacionados el uno con el otro de manera inversa. Una teoría de la naturaleza de las creencias a duras penas puede dejar fuera una característica importante como esta. No podemos tratar ‘creencias-de-que-el-césped-es-verde’, ‘creencias-de-que-el-césped-es-azul’, etc. como predicados primitivos no relacionados. Por lo tanto necesitaremos una teoría más sofisticada, una que involucre algún tipo de aparato combinatorio. Los candidatos más prometedores son aquellos que tratan a las creencias como una relación. ¿Pero una relación con qué? Hay dos distintas cuestiones aquí. Una cuestión es cómo plantear la teoría funcional de un modo detallado. Véase Loar (1981) y Schiffer (1987) para una sugerencia en cuanto a correspondencia entre las relaciones lógicas de oraciones y las relaciones inferenciales de estados mentales. Una segunda cuestión es qué tipos de estados podrían posiblemente realizar los estados de actitudes proposicionales relacionales. Field (1978) y Fodor (en Block 1980) argumentan que para explicar la productividad de estados de actitudes proposicionales no hay más alternativa que postular un lenguaje del pensamiento, un sistema de objetos estructurados sintácticamente en el cerebro que expresan las proposiciones en actitudes proposicionales. Véase Stalnaker (1984), capítulos 1-3 para una crítica del enfoque de Field. En un trabajo posterior, Fodor (1987) ha remarcado la sistematicidad de las actitudes proposicionales mencionadas arriba. Fodor señala que las creencias cuyos contenidos están relacionados sistemáticamente presentan el siguiente tipo de relación empírica: si se es capaz de creer que María ama a Juan, se es también capaz de creer que Juan ama a María. Fodor argumenta que únicamente un lenguaje del pensamiento en el cerebro podría explicar este hecho.

Externismo. El resultado de los famosos argumentos de la “Tierra Gemela” ha sido que el significado y el contenido están en parte en el mundo y en parte en la comunidad lingüística. Los funcionalistas han respondido de diversas maneras. Una respuesta es pensar los *inputs* y

outputs de una teoría funcional como de *brazo-largo* [*long-arm*] incluyendo los objetos que se ven y manipulan. Otra respuesta es adherirse a los *inputs* y *outputs* de *brazo-corto* [*short-arm*] que se detienen en la superficie del cuerpo, pensando los contenidos intencionales así caracterizados como estrechos –superviniendo a las propiedades físicas no-relacionales del cuerpo–. No ha habido una explicación ampliamente reconocida acerca de lo que es el contenido estrecho, como tampoco hay acuerdo acerca de si hay alguna carga de la prueba sobre los defensores del contenido estrecho para caracterizarlo. Véase los artículos de Burge, Loar y Stalnaker en Rosenthal (1991); véase también Goldman (1993).

Significado. El funcionalismo dice que la comprensión del significado de la palabra ‘momento’ es un estado funcional. En una versión de la perspectiva, el estado funcional puede ser visto en términos del rol de la misma palabra ‘momento’ en el pensamiento, en la solución de problemas, en la planificación, etc. Pero si la comprensión del significado de ‘momento’ es tener cierta función de la palabra, entonces hay una relación muy cercana entre el significado de una palabra y su función, y una propuesta natural es considerar la relación cercana como simple identidad, esto es, el significado de la palabra sólo *es* esa función (Véase Peacocke, 1992). Entonces el funcionalismo acerca del contenido conduce al funcionalismo acerca del significado, una teoría que pretende decirnos la naturaleza metafísica del significado. Esta teoría es popular en la ciencia cognitiva, donde en una de sus versiones es a menudo conocida como semántica procedimental [*procedural semantics*], como también en filosofía donde es a menudo conocida como semántica del rol conceptual [*conceptual role semantics*]. La teoría ha sido criticada (junto con otras versiones del funcionalismo) en Putnam (1988), y Fodor y LePore (1992).

Holismo. Block y Fodor (en Block, 1980) apuntaron el problema “carajo-caramba” [*damn/darn problem*]. Las teorías funcionales deben hacer referencia a cualquier diferencia en los estímulos o respuestas que pueden ser mentalmente significantes. La diferencia entre decir ‘carajo’ y ‘caramba’ cuando te golpeas el tobillo puede, en algunas circunstancias, ser mentalmente significativa. Por lo tanto las diferentes teorías funcionalizadas apropiadas para las dos respuestas afectarán la individuación de cada estado conectado con esas exclamaciones, y por la misma razón, cada estado conectado con esos estados, y así sucesivamente. Tus dolores conducen a ‘caramba’, los míos a ‘carajo’, por lo tanto nuestros dolores son funcionalmente diferentes y de la misma manera nuestros deseos de evitar el dolor, nuestras creencias que interactúan con esos deseos y así sucesivamente. Suposiciones plausibles conducen a la conclusión de que dos individuos que difieren de este modo no comparten casi nada en cuanto a estados mentales. El resultado es que el funcionalista necesita un modo de individuar estados mentales que sea menos refinado que apelar a una teoría completa, una caracterización molecular. Incluso si se es optimista en cuanto a la solución de este problema en el caso del dolor por medio del hallazgo de algo funcional común a todos los dolores, no puede suponerse que ese éxito se transferirá a las creencias o a los significados, mientras el éxito en el caso del significado y la creencia pueda involucrar una distinción analítico-sintético (Fodor y LePore, 1992).

Qualia. Recuérdese el autómata detector de paridad descrito al comienzo de este artículo. Éste podría ser instanciado por dos personas, cada una de las cuales está a cargo solamente de la

función especificada por una casilla. De manera similar, la organización funcional mucho más compleja de una mente humana podría “en principio” ser instanciada por medio de un vasto ejército de personas. Tendríamos que pensar al ejército como conectado a un cuerpo de robot, actuando como el cerebro de ese cuerpo y el cuerpo sería como una persona en sus reacciones a los *inputs*. ¿Pero podría tal ejército realmente instanciar una mente? Más puntualmente, ¿Podría tal ejército tener dolor o la experiencia del rojo? Si el funcionalismo atribuye mentes a cosas que no las tiene, es demasiado liberal. Lycan (1987) sugiere que incluimos mucho de psicología humana en nuestra teoría para estar funcionalizados a evitar el liberalismo; esto es, la teoría T en la definición descrita anteriormente sería una teoría psicológica más una teoría psicológica. Pero esto crea el problema opuesto, chauvinismo, peor aún. La descripción funcional resultante no se aplicará a marcianos inteligentes cuyas psicologías sean diferentes de las nuestras. Además, parece fácil imaginar un simple organismo sintiente-de-dolor que comparta un poco con nosotros lo concerniente a la organización funcional. La teoría psicológica funcionalizada de este organismo será completamente diferente de la teoría correspondiente de nosotros. Por cierto, incluso si no se adopta la táctica de Lycan, no es claro cómo el dolor podría ser caracterizado funcionalmente como para ser común a nosotros y al organismo simple (Véase mi “Troubles with Functionalism” que aparece en todas las antologías de la bibliografía).

Gran parte de la fuerza de los problemas recién mencionados derivan de la atención a estados fenoménicos tales como ver rojo. Las propiedades fenoménicas parecerían intrínsecas a (propiedades no-relacionales de) los estados que las tienen y por lo tanto las propiedades fenoménicas parecen independientes de las relaciones entre los estados, los *inputs* y los *output* que definen a los estados funcionales. Considérese, por ejemplo, el hecho de que los pacientes [a los que se les ha practicado una] lobotomía a menudo dicen seguir sintiendo dolores al igual que antes, pero que esos dolores no los molestan. Si el concepto de dolor es un concepto funcional, lo que estos pacientes dicen es contradictorio o incoherente –pero para muchos de nosotros parece ser inteligible–. Todas las antologías tienen artículos acerca de este tema; véase también Lycan (1987), capítulos 8, 9, 14 y 15 de Shoemaker (1984) y Hill (1991).

El problema chauvinismo-liberalismo afecta la caracterización de los *inputs* y los *outputs*. Si caracterizamos los *inputs* y los *outputs* de un modo apropiado para nuestros cuerpos, excluimos chauvinistamente criaturas cuya conexión con el mundo es diferente de la nuestra, v. gr. criaturas cuyas extremidades son ruedas, o para hacer la diferencia aún mayor, criaturas gaseosas que pueden manipular y percibir gases pero para los cuales todos los sólidos y los líquidos son iguales. La alternativa obvia de caracterizar los *inputs* y *outputs* en sí mismos de manera funcional parecería dar lugar a una estructura abstracta que puede ser satisfecha por, v. gr. la economía de Bolivia bajo la manipulación de un excéntrico acaudalado y caería de ese modo en el problema opuesto del liberalismo.

Es tentador responder al problema del chauvinismo suponiendo que la misma teoría funcional que se aplica a mí también se aplica a criaturas con ruedas. Si ellos piensan que tienen pies, tratarían de actuar como nosotros y si nosotros pensamos que tenemos ruedas, trataríamos de actuar como ellos. Pero nótese que las definiciones funcionales tienen que tener algunas especificaciones de órganos de *output* en ellas. Ser neutral entre todos los tipos de cuerpo que

un ser sintiente podría tener, sería adoptar la alternativa liberal de especificar los *inputs* y *outputs* en sí mismos de manera funcional. Algunos suponen que el problema puede ser manejado por medio de *outputs* condicionales. Por ejemplo, querer llevar la pelota hasta el límite del campo de juego puede ser definido en parte por medio de la tendencia a patearla si es que se tiene extremidades [*limbs*] de un cierto tipo, empujarla si se tiene extremidades de otro tipo, etc. Pero no es seguro que el ‘etc.’ pueda ser alguna vez completado en la medida que requeriría enumerar y describir físicamente cada tipo de extremidad [*appendage*] que cada tipo de ser inteligente pudiera tener. Además, el resultado de determinado deseo sobre los propios miembros depende también de cómo estén relacionados con su forma física. En el contexto de una conexión “incorrecta”, el deseo de llevar la pelota hasta el límite del campo de juego ocasionaría que la pelota se clave en la propia boca en vez de ser impulsada por el campo de juego. Pero esto hace parecer que el problema requerirá que vayamos mas lejos de lo que podría decirse que está implícito en el sentido común.

Teleología. Muchos filósofos (véase los artículos de Lycan y Sober en Lycan (1990), y Lycan (1987)) proponen que evitamos el liberalismo caracterizando los roles funcionales de manera teleológica. Excluimos los ejércitos y las economías mencionadas porque sus estados no son *para* las cosas correctas. Un problema importante para este punto de vista es la falta de una explicación teleológica aceptable. Las explicaciones basadas en la evolución chocan contra el problema de los abuelos-pantano [*swamp-grandparents problem*]. Supóngase que descubre que sus abuelos fueron formados a partir de partículas de un pantano que llegaron a unirse por azar. Por lo tanto, como sucedió esto, usted no posee ninguna historia evolutiva que contar. Si la explicación evolutiva de los apuntalamientos teleológicos del contenido son correctos, [entonces] sus estados no tienen ningún contenido. Una teoría con semejante consecuencia debería ser rechazada.

Causación. El funcionalismo establece que las propiedades mentales son propiedades de segundo orden, propiedades que consisten en tener otras propiedades que tienen ciertas relaciones entre sí. Pero hay por lo menos *prima facie* un problema en cuanto a cómo tales propiedades de segundo orden podrían ser causales y explicativas de un modo apropiado para lo mental. Considérese, por ejemplo, la provocatividad [*provocativeness*], la propiedad de segundo orden que consiste en tener alguna propiedad de primer orden (digamos rojedad [*redness*]) que causa que los toros estén enfadados. La rojedad de la capa provoca al toro, pero ¿la provocatividad de la capa provoca al toro? La provocatividad de la capa podría provocar a una sociedad protectora de animales, ¿pero no es el toro demasiado estúpido como para ser provocado por ésta? Véase Block (1990).

El funcionalismo continúa siendo un vívido y fluido punto de vista. Desarrollos positivos en los años recientes incluyen perspectivas mejoradas para el funcionalismo conceptual y la articulación del punto de vista teleológico. Desarrollos críticos incluyen problemas con la causalidad y el holismo, y continúan la controversia acerca del chauvinismo y el liberalismo.

Traducción: Joel J. Lorenzatti

Versión de revisión: 0.6