
NED BLOCK, “EL FUNCIONALISMO”

Título original: “Functionalism”. En: D. Borchert (ed.) (1996). *The Encyclopedia of Philosophy: Supplement*. New York, NY: Macmillan.

Traducción: Joel. J. Lorenzatti

¿Qué es el funcionalismo? El funcionalismo es una de las principales propuestas ofrecidas como solución al problema mente-cuerpo. Las soluciones al problema mente-cuerpo normalmente tratan de dar respuesta a cuestiones como: ¿Cuál es la naturaleza última de lo mental? En términos más generales, ¿qué es lo que hace que un estado mental sea mental? O, más específicamente, ¿qué tienen en común los pensamientos en virtud de lo cual son pensamientos? Es decir, ¿qué es lo que hace que un pensamiento sea un pensamiento? ¿Qué es lo que hace que un dolor sea un dolor? El dualismo cartesiano afirmaba que la naturaleza última de lo mental debía encontrarse en una sustancia mental especial. El conductismo identificó los estados mentales con disposiciones conductuales, mientras que el fisicismo, en su versión más influyente, identifica los estados mentales con estados cerebrales¹. El funcionalismo afirma que los estados mentales están constituidos tanto por las relaciones causales que mantienen entre sí, como por las relaciones causales que mantienen con los *inputs* sensoriales y los *outputs* conductuales². El funcionalismo constituye uno de los desarrollos teóricos más importantes de la

1N. del 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’. Al provenir del adjetivo ‘físico’, la derivación inmediata al español debió ser ‘fiscicismo’, siguiendo el mismo criterio que para la derivación inglesa de ‘physical’ nos encontramos con ‘physicalism’. El mismo criterio adopto para el término inglés ‘externalism’, el cual traduzco por ‘externismo’ y no por ‘externalismo’. No ocurre lo mismo con el término ‘funcionalismo’, puesto que al derivarse del adjetivo ‘funcional’, acaba por convertirse en ‘funcionalismo’ y no en ‘funcionismo’.

2 N. del T.: Se trata de las *entradas sensoriales* y las *salidas conductuales*. Utilizo la terminología de *inputs* y *outputs* en concordancia con la traducción de Eleonora Baringoltz en Block (1995). Las dificultades del funcionalismo (selección). En E. Rabossi (comp.). *Filosofía de la mente y ciencia cognitiva* (pp. 105-142). Barcelona, España: Paidós.

filosofía analítica del siglo XX y proporciona el fundamento conceptual de buena parte del trabajo que se hace en el área de las ciencias cognitivas.

El funcionalismo cuenta con tres fuentes distintas. En primer lugar, Putnam y Fodor entendieron los estados mentales en términos de una teoría empírica y computacional de la mente. En segundo lugar, los análisis “tópicamente neutrales” de Smart condujeron a Armstrong y Lewis a un análisis funcionalista de los conceptos mentales. En tercer lugar, la idea de Wittgenstein del significado como uso dio lugar a una versión de funcionalismo como teoría del significado, desarrollada por Sellars y, posteriormente, por Harman.

Puede apreciarse una de las motivaciones que está por detrás del funcionalismo al considerar conceptos de artefactos como *carburador* o conceptos biológicos como *riñón*. Lo que hace que algo sea un carburador es su capacidad de 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 su rol en la filtración de la sangre y la conservación de ciertos balances químicos—.

El tipo de función relevante en el caso de la mente se puede representar a partir del autómata detector de paridad, que se ilustra en la Figura 1, más abajo; el cual revela si ha visto un número par o impar de ‘1’s. Este autómata tiene dos estados: S_1 y S_2 ; dos *inputs*: ‘1’ y ‘0’; y dos *outputs*: emite la palabra “par” o “impar”. La tabla describe dos funciones: una que va desde el *input* y el estado hasta el *output*, y otra que va desde el *input* y el estado hasta el próximo estado. Cada casilla codifica dos condicionales que especifican el *output* y el próximo estado, dados el estado actual y el *input*. La casilla superior 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’s) y pasa a 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 . La casilla inferior izquierda dice que si la máquina está en S_1 y ve un ‘0’, dice “impar” y permanece en S_1 . Se prevé que la máquina comience en S_1 , de modo que si su primer *input* es un ‘0’, dirá erróneamente que vio un número impar de ‘1’s, pero una vez que haya visto uno, las respuestas siguientes serán correctas. (La falla se corrige en la siguiente máquina).

La máquina de la Figura 2 es más simple. Igual que antes, este autómata tiene dos estados: S_1 y S_2 ; y dos *outputs*: “par” e “impar”. La diferencia reside en que sólo tiene un *input*: ‘1’, aunque, por supuesto, podría no recibir ningún *input* (como la máquina de la

Figura 1). Al igual que antes, la tabla describe dos funciones: una que va desde el *input* y el estado hasta el *output*, y otra que va desde el *input* y el estado hasta el próximo estado. Al igual que antes, cada casilla codifica dos condicionales que especifican 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 vio un número impar de ‘1’s) y pasa a 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 . Esta máquina es más simple que la de la Figura 1, pero intuitivamente sirve para los mismos propósitos, al tiempo que evita señalar la ausencia de ‘1’s como un número impar de ‘1’s.

	S_1	S_2
1	“Impar” S_2	“Par” S_1
0	“Impar” S_1	“Par” S_2

Figura 1. Autómata de paridad con dos *inputs*.

	S_1	S_2
1	“Impar” S_2	“Par” S_1

Figura 2. Autómata de paridad con un solo *input*.

Supongamos ahora que preguntamos: “¿Qué es S_1 ?”. La respuesta será que la naturaleza de S_1 es completamente relacional y que se encuentra completamente capturada por la tabla. Podríamos dar una caracterización explícita de ‘ S_1 ’ (según la Figura 2) de la siguiente manera:

Estar en S_1 = Estar en el primero de dos estados que se encuentran tanto relacionados entre sí, como con los *inputs* y *outputs* de la siguiente manera: estar en alguno de los estados y recibir un *input* de ‘1’ hace que se pase al segundo estado y se

emita “impar”; y estar en el segundo de los dos estados y recibir un *input* de ‘1’ hace que se pase al primer estado y se emita “par”.

Para hacer la cuantificación sobre los estados más explícita:

Estar en $S_1 =$ Ser un x tal que $\exists P \exists Q$ [si x está en P y recibe un *input* de ‘1’, entonces pasa a Q y emite “impar”; si x está en Q y recibe un *input* ‘1’, pasa a P y emite “par” & x está en P] (Nota: léase ‘ $\exists P$ ’ como ‘hay una propiedad P ’).

Partiendo de este ejemplo, se pueden hacer algunas observaciones. (1) De acuerdo con el funcionalismo, la naturaleza de un estado mental es la misma que la del estado de un autómata: está constituida por sus relaciones con otros estados y con los *inputs* y *outputs*. Ser S_1 no es más que estar en un estado tal que al recibir un *input* de ‘1’ suceda tal o cual cosa. De acuerdo con el funcionalismo, tener un dolor no es más que tener una disposición a decir ‘ay’, preguntarse si se está enfermo, desviar la atención, etc. (2) En la medida en que los estados mentales son como los estados de un autómata, el método ejemplificado para definir los estados de autómata también debería funcionar para definir los estados mentales. Los estados mentales pueden ser totalmente caracterizados en términos que involucren sólo un lenguaje lógico-matemático, términos para señales de *input* y términos para *outputs* conductuales. De esta manera, 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: propiedades mecánicas, hidráulicas o electrónicas; pongamos por caso, que tienen ciertas relaciones entre sí. De estas propiedades, cuantificadas sobre las definiciones recién dadas, se dice que son *realizaciones* de las propiedades funcionales³. 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, algo que no habría agradado a los conductistas. (4) Un estado funcional puede ser realizado de diferentes maneras. Por ejemplo, una máquina de metal y plástico que satisfaga la tabla de la máquina podría estar hecha con engranajes, ruedas, poleas y otras cosas por el estilo —y, en tal caso, la realización de S_1 sería un estado mecánico—; o la realización de S_1 podría ser un estado electrónico, etc. (5) De la misma manera que un estado funcional

3N. del T.: El verbo ‘to realize’ y el sustantivo ‘realization’ son en este contexto términos técnicos que traduzco por ‘realizar y ‘realización’ respectivamente. Se dice así que una determinada propiedad física puede instanciar, implementar o *realizar* una propiedad mental, o que constituye una instanciación, implementación o *realización* de una propiedad mental.

puede ser realizado de diferentes maneras, un estado físico podría realizar diferentes estados funcionales en diferentes máquinas. Esto podría suceder, por ejemplo, en el caso de que se use un mismo tipo de transistor para hacer diferentes cosas en máquinas diferentes. (6) Dado que S_1 puede ser realizado de diversas maneras, la afirmación de que S_1 es un estado mecánico sería falsa, al igual que la afirmación de que S_1 es un estado electrónico. Por este motivo, existen fuertes razones para pensar que el funcionalismo muestra que el fisicismo es falso: si una criatura sin cerebro puede pensar, el pensamiento no puede ser un estado cerebral. (No obstante, véase la sección sobre funcionalismo y fisicismo más abajo).

La noción de realización amerita una discusión adicional. En los inicios del funcionalismo, se consideraba por lo general que una propiedad de primer orden realizaba una propiedad funcional en virtud de una correspondencia de uno a uno entre los dos tipos de propiedades. Pero una definición de realización de este tipo produce demasiadas realizaciones. Supóngase, por ejemplo, que en t_1 gritamos ‘uno’ frente a un balde de agua y luego en t_2 gritamos nuevamente ‘uno’. Nada nos impediría considerar al balde como un autómatas detector de paridad haciendo corresponder la configuración física del balde en t_1 con S_1 y al 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 al calor intercambiado con el medio en t_2 con “par”, y así sucesivamente. No obstante, lo que esta manera *post hoc* de pensar la correlación de la realización deja afuera, es que una verdadera realización debería satisfacer los *contrafácticos* implícitos en la tabla. Para ser una realización de S_1 , no es suficiente con que ésta conduzca a un cierto *output* y estado, dado un *input* de ‘1’: también se requiere que si el *input* hubiese sido ‘0’, la realización de S_1 hubiese conducido al otro *output* y estado. La satisfacción de los contrafácticos relevantes está incluida en la noción de realización señalada en (3). Véase Lycan (1987).

Supóngase que tenemos una teoría de los estados mentales que especifica todas las relaciones causales entre los estados, los *inputs* sensoriales y los *outputs* conductuales. Tomando al dolor como ejemplo de estado mental, podría decirse que sentarse sobre una tachuela causa dolor y que el dolor causa malestar y la exclamación ‘ay’. Si aceptamos, por mor del ejemplo, seguir adelante con esta torpe teoría, el funcionalismo diría que podemos definir ‘dolor’ de la siguiente manera: tener dolor = estar en el primero de dos

estados, el cual es causado por sentirse sobre tachuelas y que a su vez causa el otro estado y la exclamación ‘ay’. Dicho de 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 exclamación ‘ay’ & x está en P]

Dicho de manera más general, si T es una teoría psicológica con n términos mentales, de los cuales el 17^{mo} es ‘dolor’, podemos definir ‘dolor’ en relación con T de la siguiente manera (‘F₁’... ‘F_n’ son variables que reemplazan los términos mentales n, e i₁, 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 está 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 éstos. Así, el funcionalismo caracteriza lo mental en términos de estructuras que están ancladas a la realidad sólo a través de los *inputs* y *outputs*.

La teoría psicológica T recién mencionada puede ser tanto una teoría psicológica empírica como una teoría “popular” de sentido común, siendo muy diferentes los funcionalismos resultantes. En el último caso, el del *funcionalismo conceptual*, las definiciones funcionales tienen por objetivo capturar los conceptos mentales que utilizamos corrientemente. En el primer caso, que he denominado *psicofuncionalismo*, no se supone que las definiciones funcionales capturen los conceptos corrientes, sino únicamente 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 “organizativo”, análogo a la estructura computacional. El funcionalismo conceptual, por el contrario, puede pensarse como un desarrollo del conductismo lógico. Los conductistas lógicos pensaban que el dolor era una disposición a tener una *conducta* de dolor. Pero como han señalado Geach y Chisholm, lo que cuenta como conducta de dolor depende de las creencias y los deseos del agente. Los funcionalistas conceptuales evitan este problema definiendo cada estado mental en términos de su contribución a tener ciertas disposiciones de conducta y a tener otros estados mentales.

Funcionalismo y fisicismo. Con anterioridad al funcionalismo, las teorías de la mente se ocupaban de (1) qué es lo que *hay* y (2) qué es lo que da a cada tipo de estado mental su identidad propia, por ejemplo: qué es lo que tienen en común los dolores, en virtud de lo cual son dolores. Desplegando un poco estos conceptos, podemos decir que (1) es una cuestión ontológica, mientras que (2) es una cuestión metafísica. Afirmaciones ontológicas: el dualismo nos decía que existen tanto sustancias mentales como físicas, mientras que el conductismo y el fisicismo son monistas, al señalar que únicamente existen sustancias físicas. Afirmaciones metafísicas: el conductismo nos dice, por ejemplo, que lo que los dolores tienen en común, en virtud de lo cual son dolores, es algo conductual. El dualismo nos daba una respuesta no física a esta cuestión, mientras que el fisicismo nos da una respuesta física. Volviendo al funcionalismo, éste responde a la cuestión metafísica sin responder a la cuestión ontológica. El funcionalismo nos dice que lo que tienen en común los dolores —aquello que los hace ser dolores— es su función, pero no nos dice si los seres que tienen dolores poseen alguna parte no física. Esta particularidad puede ser vista en términos del autómata descrito anteriormente. Para ser un autómata como el descrito, una máquina real sólo necesita tener estados que se relacionen entre sí y con los *inputs* y los *outputs* del modo indicado. La descripción de la máquina no nos dice cómo opera o de qué está hecha y, en particular, tampoco descarta una máquina operada por un alma inmaterial, siempre y cuando esta alma pueda operar del modo determinista que especifica la tabla. Véase Putnam, *op. cit.* y el artículo de Fodor en Block (1980).

A la hora de pensar la relación entre funcionalismo y fisicismo, resulta útil distinguir dos tipos de tesis fisicistas: una versión del fisicismo se encuentra en competencia con el funcionalismo al hacer afirmaciones metafísicas respecto de la naturaleza física de las propiedades o los tipos de estados mentales (y es, por consiguiente, con frecuencia denominado fisicismo de “tipos”). Como se mencionaba anteriormente, desde un determinado punto de vista, el funcionalismo muestra que el fisicismo de tipos es falso.

Sin embargo, hay formas de fisicismo más modestas, cuyas motivaciones no son metafísicas sino ontológicas. Las afirmaciones fisicistas de este tipo no son en absoluto incompatibles con el funcionalismo. Considérese, por ejemplo, una forma de fisicismo que dice que toda cosa real está compuesta en su totalidad por partículas como las que

componen a la materia inorgánica. Tomado en este sentido, la mayoría de los funcionalistas han sido fisicistas. El funcionalismo incluso podría ser modificado en la dirección del fisicismo, por ejemplo, requiriendo que todas las propiedades cuantificadas en una definición funcional sean propiedades físicas. El fisicismo de tipos frecuentemente se contrasta con el fisicismo de *casos*. (La palabra ‘leer’ en esta oración tiene cuatro casos de letra pero sólo tres tipos de letra). El fisicismo de casos dice que cada dolor (por ejemplo) es un estado físico, pero deja abierta la posibilidad de que no haya nada físico que tengan en común todos los dolores, nada físico que haga que un dolor sea un dolor.

Es una peculiaridad de la literatura del funcionalismo y del fisicismo el hecho de que, mientras algunos funcionalistas dicen 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 dicen 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, la cuestión es en parte terminológica. Lewis es un funcionalista conceptual en cuanto a lo que es *tener dolor*. En la reglamentación de Lewis ‘tener dolor’ podría decirse que es un designador rígido de una propiedad funcional. [Un designador rígido nombra la misma cosa en todos los mundos posibles. ‘El color del cielo’ es no rígido puesto que nombra al color rojo en mundos donde el cielo es rojo. ‘Azul’ es rígido puesto que nombra al color azul incluso en mundos donde el cielo es rojo]. ‘Dolor’, por el contrario, es un designador no rígido que equivale conceptualmente a una descripción definida de la forma ‘el estado con tal y cual rol causal’. El referente de esta frase en nosotros, sostiene Lewis, es un determinado estado cerebral mientras que el referente de esta frase en un robot podría ser un estado de sus circuitos y el referente en un ángel podría ser un estado no físico. De manera similar, ‘el número ganador’ designa a ‘17’ en una lotería y a ‘596’ en otra. Por lo tanto, Lewis es funcionalista (un funcionalista conceptual, por cierto) en cuanto a lo que es tener dolor. Respecto de la cuestión metafísica descrita anteriormente —qué tienen en común los dolores en virtud de lo cual son dolores—, Lewis es funcionalista, no fisicista. Lo que tienen en común mis dolores con los de un robot es un rol causal, no algo físico. Así como no existe una similitud numérica entre 17 y 596 en virtud de la cual sean números ganadores, no existe una similitud física entre los dolores humanos y los dolores de un marciano en virtud de la cual sean dolores —como tampoco existe algún tipo de similitud física entre los dolores humanos y los dolores de los ángeles—. Sin embargo, en cuanto

a la cuestión de la naturaleza científica del dolor, Lewis es fisicista. Lo que tienen en común el dolor humano y el marciano, desde su perspectiva, no es algo científico sino conceptual.

Funcionalismo y actitudes proposicionales. La discusión respecto de la caracterización funcional ofrecida anteriormente 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., una opción teórica parecería consistir en listar simplemente los estados y sus relaciones con otros estados, *inputs* y *outputs*. Pero por varias razones, esto no constituye una opción teórica apropiada para estados de creencia, estados de deseo y otros estados de actitud proposicional. Para empezar, la lista sería demasiado larga como para realizarla sin el auxilio de métodos combinatorios. De hecho, no existe un límite en cuanto al número de proposiciones que podrían, en principio, ser un objeto del pensamiento. Por otro lado, entre las creencias existen relaciones sistemáticas, por ejemplo, entre la creencia de que Juan ama a María y la creencia de que María ama a Juan. Estos estados de creencia representan los mismos objetos pero relacionados entre sí de manera inversa. Una teoría de la naturaleza de las creencias no podría pasar por alto una característica tan importante como ésta. No podemos considerar que ‘creer-que-el-césped-es-verde’, ‘creer-que-el-césped-es-azul’, etc., son predicados primitivos que no se encuentran relacionados. De modo que necesitamos una teoría más sofisticada que involucre algún tipo de aparato combinatorio. Las candidatas más prometedoras son aquellas que consideran que las creencias son una relación. ¿Pero relación con qué? Existen aquí dos cuestiones distintas. Una tiene que ver con cómo plantear la teoría funcional de una manera detallada. Véase Loar (1981) y Schiffer (1987) para una sugerencia en cuanto a la correspondencia entre las relaciones lógicas entre oraciones y las relaciones inferenciales entre estados mentales. La otra cuestión tiene que ver con qué tipos de estados podrían llegar a realizar a los estados de actitud proposicional que son relacionales. Field (1978) y Fodor (en Block 1980) sostienen que, para explicar la productividad de los estados de actitud proposicional, no hay otra alternativa que postular un lenguaje del pensamiento, un sistema de objetos estructurados sintácticamente en el cerebro que expresen las proposiciones en actitudes proposicionales. Véase Stalnaker (1984), capítulos 1-3 para una crítica del enfoque de Field. En un trabajo posterior, Fodor (1987) acentuó la sistematicidad propia de las actitudes proposicionales que hemos mencionado recién. Fodor señala que las creencias cuyos contenidos están

relacionados de manera sistemática presentan relaciones empíricas del siguiente tipo: si somos capaces de creer que María ama a Juan, también somos capaces de creer que Juan ama a María. Fodor sostiene que sólo un lenguaje del pensamiento localizado en el cerebro podría explicar este hecho.

Externismo. La conclusión de los famosos argumentos de la “Tierra Gemela” fue que significado y contenido están en parte en el mundo y en parte en la comunidad lingüística. Los funcionalistas respondieron de diversas maneras. Una respuesta ha sido pensar los *inputs* y *outputs* de una teoría funcional como de *brazo largo*, incluyendo a los objetos que uno puede ver y manipular. Otra respuesta ha sido quedarse con *inputs* y *outputs* de *brazo corto*, deteniéndose en las superficies del cuerpo, concibiendo los contenidos intencionales como *estrechos* —superviniendo a las propiedades físicas no relacionales del cuerpo—. No existe aún una explicación ampliamente reconocida sobre qué pueda ser el contenido estrecho, así como tampoco hay consenso respecto de si pesa algún tipo de carga de la prueba sobre los defensores del contenido estrecho a la hora de caracterizarlo. Véase los artículos de Burge, Loar y Stalnaker en Rosenthal (1991) y véase también Goldman (1993).

Significado. El funcionalismo dice que comprender el significado de la palabra ‘momento’ es un estado funcional. Desde una de las versiones de esa perspectiva, ese estado funcional puede verse en términos del rol de la palabra ‘momento’ en el pensamiento, en la solución de problemas, en la planificación, etc. Pero si comprender el significado de ‘momento’ reside en que la palabra tenga cierta función, entonces existe una estrecha relación entre el significado de una palabra y su función; y una de las propuestas naturales sería considerar esta relación simplemente como una identidad, es decir, el significado de la palabra vendría a *ser* esa función (Véase Peacocke 1992). De esta manera, el funcionalismo sobre el contenido conduce al funcionalismo sobre el significado, una teoría que pretende revelarnos la naturaleza metafísica del significado. Esta teoría es popular tanto en las ciencias cognitivas, donde una de sus versiones es frecuentemente conocida como semántica procedimental, como en filosofía, donde es frecuentemente conocida como semántica del rol conceptual. Esta teoría fue criticada (junto con otras versiones del funcionalismo) en Putnam (1988) y en Fodor y LePore (1992).

Holismo. Block y Fodor (en Block 1980) advirtieron el problema “carajo-caramba”. Las teorías funcionales deben hacer referencia a alguna diferencia en los estímulos o en las respuestas que puedan ser mentalmente significativas. La diferencia entre decir ‘carajo’ y decir ‘caramba’ cuando uno se golpea el tobillo puede, en determinadas circunstancias, ser mentalmente significativa. Por lo tanto, las diferentes teorías funcionalizadas apropiadas para cada respuesta afectarán la individuación de cada estado relacionado con esas exclamaciones y, por la misma razón, a cada estado relacionado con esos estados, y así sucesivamente. Los dolores que a uno le hacen exclamar ‘caramba’, a otros le hacen exclamar ‘carajo’, por lo tanto, son dolores funcionalmente diferentes. Lo mismo ocurre con los respectivos deseos de evitar el dolor, las creencias ligadas a esos deseos, y así sucesivamente. Suposiciones plausibles nos llevan a concluir que dos individuos que difieran de esta manera prácticamente no tendrían estados mentales en común. La conclusión es que el funcionalista necesita un modo de individuar los estados mentales que sea menos detallado que la teoría en su totalidad, necesita una caracterización de tipo molecular. Pero, aun siendo optimistas en cuanto a la solución de este problema en lo que respecta al caso del dolor, es decir, encontrando algo funcional que sea común a todos los dolores, no puede suponerse que ese éxito se haga extensivo a las creencias o a los significados: el éxito en el caso de los significados y las creencias podría involucrar una distinción analítico-sintético (Fodor y LePore, 1992).

Qualia. Recuérdese el autómata detector de paridad descrito al comienzo del artículo. El autómata podría ser instanciado por dos personas, cada una a cargo de la función especificada por una de las casillas. De manera similar, la organización funcional de una mente humana, que es mucho más compleja, podría “en principio” ser instanciada por un numeroso ejército de personas. Tendríamos que pensar en el ejército conectado a un cuerpo de robot, haciendo de cerebro, y ese cuerpo sería como una persona en lo tocante a las reacciones a los *inputs*. ¿Pero podría ese ejército, en efecto, instanciar una mente? Más puntualmente: ¿podría ese ejército tener un dolor o la experiencia del color rojo? Si el funcionalismo atribuye mente a cosas que no la tienen, es demasiado liberal. Con el fin de evitar el liberalismo, Lycan (1987) sugiere incluir más de fisiología humana en la teoría que vayamos a funcionalizar. Es decir, la teoría T, según la definición provista anteriormente, sería una teoría psicológica *más* una teoría fisiológica. Pero esto crea el problema opuesto, el del chauvinismo, que es peor. La descripción funcional resultante

no podría atribuirse a marcianos inteligentes, cuyas fisiologías sean diferentes de las nuestras. Además, parece fácil imaginar un organismo simple que sienta dolor pero que tenga poco que ver con nosotros en cuanto a su organización funcional. La teoría fisiológica funcionalizada de este organismo sería irremediabilmente diferente de la nuestra. De hecho, incluso sin adoptar la táctica de Lycan, no resulta claro cómo podría caracterizarse al dolor en términos funcionales de modo tal que pueda ser común a nosotros y al organismo simple recién mencionado (Véase mi artículo “Troubles with Functionalism”, incluido en todas las antologías de la bibliografía⁴).

Buena parte de la fuerza que tienen los problemas recién mencionados surge de considerar estados fenoménicos tales como la apariencia del rojo. Las propiedades fenoménicas pareciesen ser intrínsecas a (propiedades no relacionales de) los estados que las tienen y, por lo tanto, independientes de las relaciones entre los estados, los *inputs* y los *output* que definen a los estados funcionales. Considérese, por ejemplo, a los pacientes a los que se les ha practicado una lobotomía, los cuales frecuentemente manifiestan seguir sintiendo dolores al igual que antes, pero que esos dolores ya no les molestan. Si el concepto de dolor es un concepto funcional, lo que estos pacientes dicen es contradictorio o incoherente —aunque a muchos de nosotros nos parezca inteligible—. Todas las antologías tienen artículos acerca de este tema; véase también Lycan (1987), los capítulos 8, 9, 14 y 15 de Shoemaker (1984) y Hill (1991).

El problema del chauvinismo-liberalismo afecta la caracterización de los *inputs* y los *outputs*. Si caracterizamos los *inputs* y los *outputs* de modo tal que sea apropiado para nuestros cuerpos, excluimos chauvinistamente criaturas cuyo contacto con el mundo es diferente del nuestro, v. gr. criaturas cuyas extremidades terminan en ruedas o, para acentuar la diferencia, criaturas compuestas de gas que pueden percibir y manipular gases pero para las cuales todos los objetos sólidos y líquidos resultan iguales. La alternativa obvia, caracterizar a los *inputs* y *outputs* mismos de manera funcional, pareciera dar como resultado una estructura abstracta capaz de ser satisfecha por, v. gr., la economía de Bolivia bajo la manipulación de un excéntrico acaudalado, y por lo tanto, caería en el problema opuesto, el del liberalismo.

4 N. del T.: Hay traducción al español en E. Rabossi (comp.) (1995). *Filosofía de la mente y ciencia cognitiva*. Barcelona, España: Paidós.

Resulta tentador responder al problema del chauvinismo suponiendo que la misma teoría funcional que se aplica en mi caso también podría aplicarse en criaturas con ruedas. Si ellos creyesen tener pies tratarían de actuar como nosotros y, si nosotros creyésemos tener ruedas, trataríamos de actuar como ellos. Sin embargo, nótese que las definiciones funcionales deben incorporar ciertas especificaciones de los órganos de *output*. Permanecer neutrales respecto de todos los tipos de cuerpo que los seres con sensaciones podrían tener sería adoptar la alternativa liberal de especificar los *inputs* y *outputs* mismos de manera funcional. Algunos imaginan que el problema puede manejarse apelando a *outputs* condicionales. Querer llevar la pelota hasta el extremo del campo de juego, por ejemplo, podría, en parte, definirse como la tendencia a patear la pelota si se tiene extremidades de un cierto tipo, a empujarla si se tiene extremidades de otro tipo, etc. Sin embargo, no resulta claro que tal ‘etc.’ pueda siquiera completarse alguna vez, desde el momento en que esto supondría enumerar y describir físicamente todos los tipos de extremidades que pudiera llegar a tener un ser inteligente. Además, el resultado de un determinado deseo en el uso de una extremidad depende tanto de la forma física que ésta tenga como de su modo de conexión con el cuerpo. En el caso de una conexión “incorrecta”, el deseo de llevar la pelota hasta el extremo del campo de juego podría tener como resultado que la pelota se nos atasque en la boca en lugar de ser impulsada por el campo de juego. Pero este punto hace parecer que el problema requerirá ir mucho más allá de lo que pueda decirse desde el sentido común.

Teleología. Varios filósofos (véase los artículos de Lycan y Sober en Lycan (1990), y Lycan (1987)) proponen evitar el liberalismo caracterizando los roles funcionales en términos teleológicos. Excluimos a los ejércitos y a las economías mencionadas simplemente porque sus estados no son *para* las cosas adecuadas. Un problema importante que enfrenta esta perspectiva reside en la ausencia de una explicación teleológica aceptable. Las explicaciones basadas en la evolución tropiezan con el problema de los abuelos-pantano. Imagine que descubre que sus abuelos fueron hechos a partir de partículas de un pantano que llegaron a unirse por pura casualidad. En la medida en que esto sea lo que haya sucedido, no se podría aducir ninguna historia evolutiva. Si la explicación evolutiva del andamiaje teleológico del contenido fuese correcta, los estados de usted no tendrían ningún contenido. Una teoría con semejantes implicaciones debería rechazarse.

Causación. El funcionalismo señala que las propiedades mentales son propiedades de segundo orden, propiedades que consisten en tener otras propiedades que mantienen ciertas relaciones entre sí. Sin embargo, existe, por lo menos a primera vista, un problema respecto de cómo esas propiedades de segundo orden podrían ser causales y explicativas de un modo que sea apropiado para lo mental. Considérese, por ejemplo, la provocatividad: la propiedad de segundo orden que consiste en tener alguna propiedad de primer orden (la rojez, pongamos por caso) que causa que los toros se enfurezcan. La rojez de la capa provoca al toro ¿pero puede decirse de la misma manera que 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 una perspectiva viva y fluida. Los desarrollos positivos de los últimos años incluyen mejores perspectivas para el funcionalismo conceptual y la articulación de la perspectiva teleológica. Los desarrollos críticos incluyen problemas con la causalidad y con el holismo, y la continuación de la controversia en torno al chauvinismo y al liberalismo.

Bibliografía

- Beakley, B. y Ludlow, P. (eds.) (1992). *Philosophy of Mind: Classical Problems/Contemporary Issues*. Cambridge, MA: MIT Press.
- Block, N. (ed.) (1980). *Readings in Philosophy of Psychology* (Vol. 1). Cambridge, MA: Harvard University Press.
- Block, N. (1990). Can the mind change the world? En G. Boolos (ed.). *Meaning and Method: Essays in Honor of Hilary Putnam* (pp. 137-170). Cambridge, UK: Cambridge University Press.
- Field, H. (1978). Mental representation. En *Erkenntnis* 13 (pp. 9-61).
- Fodor, J. (1987). *Psychosemantics*. Cambridge, MA: MIT Press.
- Fodor, J. y Lepore, E. (1992). *Holism*. Oxford, UK: Blackwell.

Goldman, A. (1993). *Readings in Philosophy and Cognitive Science*. Cambridge, MA: MIT Press.

Hill, Ch. (1991). *Sensations*. Cambridge, UK: Cambridge University Press.

Loar, B. (1981). *Mind and Meaning*. Cambridge, UK: Cambridge University Press.

Lewis, D. (1995). Reduction of Mind. En S. Guttenplan (ed.). *A Companion to Philosophy of Mind*. Oxford, UK: Blackwell.

Lycan, W. (1987). *Consciousness*. Cambridge, MA: MIT Press.

Lycan, W. (ed.) (1990). *Mind and Cognition*. Oxford, UK: Blackwell.

Peacocke, C. (1992). *A Study of Concepts*. Cambridge, UK: MIT Press.

Putnam, H. (1988). *Representation and Reality*. Cambridge, MA: MIT University Press.

Rosenthal, D. (ed.) (1991). *The Nature of Mind*. Oxford, UK: Oxford University Press.

Schiffer, S. (1987). *Remnants of Meaning*. Cambridge, MA: MIT Press.

Shoemaker, S. (1984). *Identity, Cause and Mind*. Ithaca: Cornell University Press.

Stalnaker, R. (1984). *Inquiry*. Cambridge, MA: MIT Press.