



UNIVERSIDAD NACIONAL DE COLOMBIA

EN BUSCA DE UNA APROXIMACIÓN FISICALISTA A LA CONCIENCIA: Modalidad, Subjetividad, Brecha Explicativa y el Problema Duro

Luis Alejandro Murillo Lara

Universidad Nacional de Colombia

Facultad de Ciencias Humanas

Departamento de Filosofía

Bogotá, Colombia

2011

EN BUSCA DE UNA APROXIMACIÓN FISICALISTA A LA CONCIENCIA: Modalidad, Subjetividad, Brecha Explicativa y el Problema Duro

Luis Alejandro Murillo Lara

Tesis presentada como requisito parcial para optar al título de:

Magister en Filosofía

Director:

Doctor Jaime Ramos Arenas

Línea de Investigación:

Filosofía de la Mente

Universidad Nacional de Colombia

Facultad de Ciencias Humanas

Departamento de Filosofía

Bogotá, Colombia

2011

A mis seres queridos, quienes han decidido iluminar mi vida con su presencia: Cristina, mi madre; Marcela, Sonia y Viviana, mis hermanas; y Lizeth, mi compañera.

Resumen

En este trabajo me he trazado dos propósitos fundamentales. En primer lugar deseo hacer una revisión crítica de algunos de los principales argumentos que fundaron el debate alrededor de estatus ontológico de la conciencia fenoménica. Me concentraré en los argumentos dirigidos a problematizar el fisicalismo en la filosofía de la mente. En segundo lugar, deseo plantear hacia dónde me parece que apuntan dichos argumentos y discusiones, qué es lo que considero que establecen y qué es lo que considero que fallan en probar; esto dará lugar a sugerir un camino por el cual podría encontrarse una salida que le permita al fisicalismo mantenerse en pie. Sospecho que la mayor parte de las objeciones contra el fisicalismo pueden ser abordadas desde el punto de vista que quiero proponer, según el cual se conciba que tenemos dos maneras de darse de una misma realidad (como experiencias conscientes ante el sujeto de dichos estados, como activaciones en una materia gris ante el observador externo de dichos estados).

Palabras clave: filosofía de la mente, conciencia, fisicalismo, experiencia.

Abstract

I have two major purposes in this work. Firstly, I want to do a critical review of some of the main arguments that founded the debate on the ontological status of phenomenal consciousness. I will focus on the arguments aimed to put into question the physicalism in the philosophy of mind. Secondly, I want to settle where I think these arguments and discussions point to, what they establish, and what they fail to prove. This will allow us to suggest a path by which the physicalism would stand. I suspect that most of the objections against physicalism can be tackled from the point of view I propose, according to which it is conceived that there are two ways in which the same reality can be given (as conscious experiences to the subject of these states, and as grey matter activations to the outer observer of the very states).

Keywords: philosophy of mind, consciousness, physicalism, experience.

Contenido

	Pág.
Resumen	VII
Introducción	1
1. Capítulo 1: TRES ARGUMENTOS CONTRA EL FISCALISMO ACERCA DE LA CONCIENCIA FENOMÉNICA.....	5
1.1 KRIKKE: CONCEBIBILIDAD Y NECESIDAD DE LA IDENTIDAD PSICOFÍSICA	5
1.1.1 El Desafío de Kripke	10
1.2 NAGEL Y LA SUBJETIVIDAD DE LA EXPERIENCIA	15
1.2.1 Subjetividad y Fisicalismo	19
1.3 JACKSON Y LA “INFORMACIÓN FÍSICA”	24
1.3.1 “INFORMACIÓN FÍSICA”	27
1.3.2 La Versión Fuerte.....	33
2. EL DEBATE ALREDEDOR DE LOS ARGUMENTOS CONTRA EL FISCALISMO .	39
2.1 LEVINE: La Brecha Explicativa.....	39
2.1.1 ¿Una <i>mera</i> cuestión epistemológica?	41
2.1.2 La Brecha Explicativa.....	42
2.1.3 El Alcance de la Brecha y el asunto de la Inteligibilidad	45
2.1.4 Dos formas de concebir el fisicalismo	51
2.2 LEVINE: FISCALISMO ONTOLÓGICO Y FISCALISMO EPISTÉMICO	53
2.2.1 La Lectura Epistemológica	58
2.3 CHALMERS: EL PROBLEMA DIFÍCIL, LA IMPOSIBILIDAD DE LA REDUCCIÓN Y EL PANPSIQUISMO.....	63
2.3.1 EL PROBLEMA DIFÍCIL Y LA IMPOSIBILIDAD DE LA REDUCCIÓN	63
2.3.2 La Dificultad de Entender el Problema Difícil.....	65
2.3.3 El Problema Difícil y la Brecha Explicativa	66
2.3.4 LA TEORÍA PANPSIQUISTA DE CHALMERS.....	70
2.3.5 La experiencia consciente como propiedad fundamental de la realidad	75
2.4 LOAR Y LOS CONCEPTOS FENOMÉNICOS.....	77
2.4.1 CONCEPTOS Y CUALIDADES: LO FÍSICO Y LO FENOMÉNICO.	78
2.4.2 LOS CONCEPTOS DE RECONOCIMIENTO.....	81
2.4.3 Las identidades <i>a posteriori</i> entre términos conceptualmente independientes.....	83
2.4.4 LA BRECHA EXPLICATIVA SEGÚN LOAR.....	85
2.4.5 La Brecha Explicativa y Argumento del Análisis Conceptual: Levine, Loar, Chalmers, Jackson, Block y Stalnaker.....	87
2.4.6 Nagel y el juego de “dejar por fuera”	90

2.4.7	Acerca de Internalismo vs Externalismo	91
2.5	PAPINEAU: TENIENDO CUIDADO CON LA BRECHA	94
2.5.1	El principal argumento a favor del fisicalismo: el argumento causal.....	94
2.5.2	¿Puede el fisicalismo salirle al paso a los reparos del funcionalista? ...	96
2.5.3	La Brecha según Papineau.....	98
2.5.4	Teniendo Más Cuidado con la Brecha	101
2.5.5	Papineau y el Rechazo de las Propiedades Fenoménicas.....	103
3.	CONCLUSIONES	105
3.1	Sobre el Desafío de Kripke.....	105
3.2	Sobre los argumentos de Nagel	106
3.3	El Argumento del Conocimiento	107
3.4	La Brecha Explicativa	108
3.5	El Hard Problem y Brecha Explicativa	109
3.6	La Premisa Semántica y la Brecha Explicativa	109
3.7	Papineau y su negación de la Brecha Explicativa.....	110
3.8	Sobre el Fisicalismo	110
3.9	Discreta delineación de una auténtica salida para fisicalista: La Teoría <i>Perspectival</i> de la Conciencia (TPC).....	113
3.9.1	La TPC y el Dualismo Conceptual	114
3.9.2	La TPC y la Brecha Explicativa.....	116
3.10	Obstáculos Potenciales para la TPC	116
3.10.1	El ' <i>Argumento del Análisis Conceptual</i> '	116
3.10.2	La ' <i>Premisa Semántica</i> '	118
3.11	Posibles Alternativas	118
3.12	Concebibilidad y Posibilidad	119
3.13	Perspectivas de Investigación Futura	121
	Bibliografía	123

Introducción

En la filosofía de la mente de las últimas décadas, cierto problema se ha convertido en una especie de tabú para los estudiosos del tema. Muestra de ello es que muchos filósofos de la mente prefieren deliberadamente volver su mirada hacia otros asuntos, mientras otros afirman con cierto desdén que se trata de un pseudoproblema y otros sostienen desalentados que no se puede resolver. En este trabajo pretendo retomar dicho problema, estimar nuevamente sus dimensiones y proponer un camino sugestivo por el cual podrían explorarse posibles salidas al mismo.

El problema en cuestión es el de la relación entre la experiencia consciente, también llamada conciencia fenoménica, y ciertos rasgos cerebrales. Durante mucho tiempo los filósofos han perdido el sueño tratando de entender cuál es el lugar de la experiencia consciente, con su subjetividad propia, en el mundo físico. Se ha propuesto que es parte de otro dominio ontológico y que no puede hacer parte del dominio de lo físico, incluso que es un mero constructo teórico dirigido a explicar la conducta, o que se trata de una función resultante de cierta organización interna que no requiere de un sustrato específico para llevarse a cabo y, por último, que es idéntica a ciertos estados cerebrales.

Por diversas razones, que iré exponiendo en su momento, pienso que el fisicalismo es el enfoque más consistente para abordar este problema. Así, en este trabajo me he trazado dos propósitos fundamentales. En primer lugar, a partir de un modesto estado del arte, deseo hacer una revisión crítica de algunos de los principales argumentos que fundaron el debate alrededor de estatus ontológico de la experiencia consciente o conciencia fenoménica. Me concentraré en los argumentos dirigidos a problematizar el fisicalismo sobre la conciencia fenoménica, y rastrearé buena parte de las discusiones que estos argumentos han suscitado. Mi intención será examinar la idea de que hay alguna clase de contradicción, o al menos una tensión entre los rasgos inherentes a la experiencia consciente y la verdad del fisicalismo. En segundo lugar, deseo plantear hacia dónde me

parece que apuntan dichos argumentos y discusiones, qué es lo que considero que establecen y qué es lo que considero que fallan en probar; esto dará lugar a sugerir un camino por el cual podría encontrarse una salida que le permita al fisicalismo mantenerse en pie.

El punto de vista que propondré involucra, a su vez, dos consideraciones: por un lado, que la relación Experiencia consciente—Ciertos procesos cerebrales tiene la forma de una identidad empírica a posteriori y, por otro lado, que existe una brecha explicativa entre los dos ítems de la identidad. Por supuesto, defender estas dos consideraciones requiere evaluar y enfrentar las principales objeciones que han recibido. Pienso que un razonamiento más o menos simple es capaz de dar cuenta de las dos consideraciones de una forma más o menos satisfactoria; esto es lo que llamaré un “enfoque perspectival”. Con la expresión “enfoque perspectival” me refiero a un tipo de análisis dentro de cual se conciba lo fenoménico (las experiencias conscientes) como una ‘perspectiva’, ‘modo de presentación’ o ‘aspecto’ de ciertos estados físicos. Tendríamos dos maneras de presentarse (de darse) de la misma cosa: una ante el sujeto de dichos estados (como experiencias conscientes, etc.), otra ante el observador de dichos estados (como actividad eléctrica en ciertas regiones de una materia gris y gelatinosa).

Como salta a la vista, la posición que planteo es una defensa del fisicalismo en la medida que de acuerdo con ella las experiencias conscientes y ciertas propiedades y procesos neurofisiológicos serían una y la misma cosa. Hace falta, sin embargo, precisar mucho más qué es lo que se va a defender, con el fin de poner al descubierto las complicaciones a las que se enfrenta el defensor de esta postura. El fisicalismo es una tesis metafísica general según la cual todo es físico¹, esto significa que es una tesis puramente ontológica, semejante a la de Tales de que todo era agua. El fisicalismo sostiene que los estados o procesos de la mente son idénticos a estados o procesos del cerebro. Por ejemplo, con respecto a las experiencias de dolor, la teoría de la identidad sostiene que estas experiencias son procesos cerebrales (y no algo meramente

¹ Obviamente los fisicalistas no niegan que el mundo pueda contener cosas que a primera vista no *parezcan* físicas (cosas de naturaleza biológica, social, etc.), pero insisten en que al fin de cuentas estos elementos son de algún modo completamente físicos.

correlacionado con procesos cerebrales). Así pues, para el fisicalista, debe ser verdad un enunciado de identidad de la forma “el dolor es idéntico a tal o cual estado/proceso cerebral”, o al menos está comprometido con la verdad de algún enunciado de una especie similar.

Durante el curso de la segunda mitad del siglo pasado, a través de investigaciones sobre la percepción, la memoria, la capacidad de discriminar y categorizar, entre otros, se admitió paulatinamente que el fisicalismo era verdad acerca de muchos rasgos de lo mental. En esta medida, en la literatura filosófica al respecto, se ha concedido de manera más o menos unánime que muchos rasgos de lo mental pueden ser de alguna manera idénticos a rasgos físicos; pero se ha mantenido un largo disenso en el que los detractores de la teoría de la identidad han alegado que existe un fenómeno de lo mental que no es físico, a saber, la conciencia. Este alegato ha sido expresado de diversas maneras: que las experiencias conscientes *tienen* propiedades no-físicas, que las experiencias conscientes *no son* estados físicos, que la conciencia en tanto fenómeno de la mente *no es* física. La idea común a todas estas formulaciones es que existe algo en la naturaleza misma de la conciencia que hace falsa la tesis fisicalista.

Para sustentar estos planteamientos, los detractores del fisicalismo han acudido a diversos argumentos que parecen mostrar que el fisicalismo acerca de la conciencia no puede ser cierto (y que por lo tanto cualquier enunciado que la identifique con algo físico tiene que ser falso). Estos argumentos pretenden denunciar problemas lógicos, epistemológicos y conceptuales que muestran la insostenibilidad del fisicalismo.

Parece haber una idea común detrás de los argumentos que han sido presentados en contra del fisicalismo: incluso cuando completara toda la descripción física del funcionamiento del cerebro, no se habría dicho nada sobre experiencias conscientes, ni se puede derivar ningún conocimiento acerca de por qué esos cambios neurobiológicos generan tal o cual experiencia consciente específica (o cualquier experiencia en general). Esto sugiere una incapacidad para ligar lo físico con la experiencia consciente y para entender cómo la conciencia podría depender exactamente de un sustrato físico.

Sospecho que la mayor parte de estas objeciones contra el fisicalismo pueden ser abordadas desde el punto de vista que quiero proponer, un punto de vista según el cual se conciba que tenemos dos maneras de darse de una misma realidad (como

experiencias conscientes ante el sujeto de dichos estados, como activaciones en una materia gris ante el observador externo de dichos estados). Como veremos, la articulación de este punto de vista seguirá una distinción entre epistemología y ontología.²

² El asunto de la conciencia fenoménica bien podría ser el tema principal de la filosofía de la mente misma: la fuente principal de nuestro conocimiento de nuestros estados mentales son nuestros estados mentales conscientes, estados fenoménicamente conscientes; prácticamente todos nuestros análisis sobre otras clases de estados mentales, o sobre rasgos mentales no-conscientes vienen de nuestro conocimiento de la existencia de estados mentales conscientes. Así, una explicación de lo mental parecería quedarse bastante corta si omite el aspecto consciente de lo mental, pues lo que esperaríamos, por supuesto, es que se procurara dar cuenta de los estados mentales que nos son más familiares, las experiencias conscientes, así como de la relación de estas con el resto de la realidad.

1. Capítulo 1: TRES ARGUMENTOS CONTRA EL FISCALISMO ACERCA DE LA CONCIENCIA FENOMÉNICA

1.1 KRIPKE: CONCEBIBILIDAD Y NECESIDAD DE LA IDENTIDAD PSICOFÍSICA

Uno de los primeros y más contundentes ataques contra el fiscalismo fue desarrollado por Saúl Kripke, y se apoya en una serie de consideraciones sobre las propiedades modales de los enunciados de identidad. En el ensayo "*Identidad y Necesidad*" (Kripke, 1972), muestra la manera en que estas consideraciones nacen de una reflexión cuyo propósito es mostrar que no existen enunciados de identidad contingentes.

Kripke empieza presentando un argumento que tiene como premisas una serie de verdades lógicas innegables:

- $\forall x \forall y [(x=y) \rightarrow (f(x)=f(y))]$

Si x y y son iguales, comparten las mismas propiedades,

- $\forall x \Upsilon(x=x)$

Todo objeto tiene la siguiente propiedad (modal): ser necesariamente igual a sí mismo,

- $\forall x \forall y [(x=y) \rightarrow (\Box(x=x) = \Box(x=y))]$

(Esta premisa es un caso de la primera premisa) Si x y y son idénticos, entonces y debe poseer todas las propiedades de x , lo cual incluye las propiedades modales de x . Así que si x tiene la propiedad modal de ser necesariamente idéntico a sí mismo, entonces y también la tiene (es decir, y tiene la propiedad de ser necesariamente idéntico a x).

- $\forall x \forall y [(x=y) \rightarrow \Box(x=y)]$

Por modus ponens entre la segunda y la tercera premisas, se sigue que si es verdad que cualesquiera x y y son idénticos, entonces es *necesario* que lo sean, pues y poseerá una propiedad de x : y posee la propiedad de ser necesariamente idéntico a x .

Con respecto a esta conclusión, Kripke hace algunas salvedades que incluyen la aclaración de que los enunciados de identidad que contienen descripciones no son necesarios, sino contingentes (“ $e! F=e! G$ ” es un hecho contingente, pues $e! F$ pudo no haber sido $e! G$).³ Kripke piensa que la contingencia de las identidades que contienen descripciones se pueden abordar sustituyendo las descripciones por cuantificadores universales: en el caso de “ $e! F=e! G$ ”, se diría que hay un objeto x y como cuestión de hecho contingente x es $e! F$ y, también como cuestión de hecho contingente, hay un objeto y que es $e! G$, y es necesario que $x=y$.

Ahora, Kripke anota que la ley de Leibniz –puesta en la primera premisa al comienzo del apartado– no vale sólo en su forma cuantificada sino también para nombres propios. Por lo tanto, el argumento del comienzo del apartado se puede repetir con nombres propios: los enunciados de identidad entre nombres propios tienen que ser necesarios si son verdaderos.

Pero la tesis de Kripke no sólo debe dar cuenta de la contingencia de los enunciados de identidad que contienen descripciones. Algunos enunciados de identidad que contienen nombres propios o términos teóricos son también *a posteriori* (“*Phosphorus es Hesperus*”, “*el calor es el movimiento de las moléculas*”, “*los estados mentales son estados físicos*”), por lo que se cree que las cosas pudieron haber sido de otra manera (pareciera que podemos imaginarnos que fuera de otra manera, que podemos concebir una situación contrafáctica o ‘mundo posible’ en el que estos enunciados son falsos: por ejemplo, puedo imaginar que mi dolor exista aún cuando el estado de mi cuerpo no existiera); y si pudo haber sido de otra manera, entonces los enunciados no son necesarios sino contingentes. Pero Kripke mantiene su posición, y señala que en todos estos casos, los enunciados de identidad son necesarios si son verdaderos (en esta medida, no hay contingencia, sólo la *ilusión* de contingencia).

Para abordar este problema, Kripke empieza introduciendo la noción de Designador Rígido. Si podemos imaginar que el mundo pudo haber sido distinto de manera que el

³ Ni siquiera el enunciado “ $e! G=e! G$ ” es necesario, es decir, $e! G$ pudo no haber sido $e! G$. ¿Cómo pudo $e! G$ no haber sido $e! G$? esto se explica acudiendo al mismo aparato russelliano de las descripciones: la descripción tiene alcance amplio en la primera ocurrencia (el que hizo tal y tal cosa tiene la propiedad de que pudo no haber hecho tal y tal cosa).

término hubiera designado a otro objeto, entonces el término no designa rígidamente. Un designador rígido es un término que designa al mismo objeto en todos los mundos posibles.⁴ Evidentemente, la expresión *designador rígido* no implica que el objeto referido tenga que existir necesariamente, sino que en cualquier mundo posible *donde exista*, usamos el designador en cuestión para designar el objeto. Así pues, de acuerdo a lo visto antes, las descripciones no son designadores rígidos, mientras que los nombres y expresiones científicas sí lo son (intuitivamente, los nombres propios parecen ser designadores rígidos: Nixon no podría no haber sido Nixon, pues de ser así, simplemente no habría existido).

El siguiente paso de Kripke es establecer que aprioricidad y necesidad no son sinónimas. Un enunciado es necesario cuando es verdadero y no podría haber sido de otra manera (este rasgo del enunciado se encuentra entonces en el ámbito de la metafísica). Un enunciado es *a priori* cuando puede saberse que es verdadero independientemente de la experiencia (nuestro conocimiento de la verdad de un enunciado es un asunto epistémico, por lo que este rasgo del enunciado hace parte de ámbito de la epistemología y no de la metafísica). Aprioricidad y necesidad son, pues, diferentes, son rasgos de los enunciados de identidad que hacen parte de ámbitos diferentes; incluso si se planteara que aprioricidad y necesidad son coextensas, bastaría pensar en enunciados cuyo valor de verdad no conocemos (es decir, cuya verdad no conocemos *a priori*), pero que bien podrían ser verdaderos (recuérdese el caso de Goldbach). Esto muestra que sólo porque un enunciado sea necesario no significa que pueda ser conocido *a priori*.⁵

De regreso a los enunciados de identidad donde aparecen nombres propios y términos científicos, Kripke señala cómo en la medida que son *a posteriori* es que se cree que pudieron haber sido de otra manera, pues pareciera que podemos imaginarnos que

⁴ En este caso hablamos del término en nuestro lenguaje, tal como *nosotros* lo usamos al especificar una situación contrafáctica.

⁵ Kripke hace aquí una alusión a la importancia que esta distinción podría tener para el esencialista. Si un objeto tiene propiedades esenciales, entonces un enunciado que afirma que el objeto tiene x propiedad esencial debe ser necesariamente verdadero, pero como no sé *a priori* cuáles son todas las propiedades de un objeto, entonces no sé *a priori* si ese enunciado es verdad.

fueran de otra manera (y si pudiera haber sido de otra manera no son necesarios sino contingentes). Pero según lo que se ha dicho, la verdad (*ergo*, la necesidad) de ciertos enunciados de identidad –donde aparecen estas expresiones– puede no ser cognoscible *a priori* sino *a posteriori*.

De nuevo, Kripke insiste en que los nombres propios y los términos científicos son designadores rígidos, por lo que las identidades son necesarias (*a* y *b* serán designadores rígidos de un cierto *x*, si se refieren a él en todo mundo posible, luego no habrá ninguna situación contrafáctica en la que *a* pudiera no haber sido *b*, porque sería una situación en la que *x* no habría sido idéntico a sí mismo). ¿Cómo se logra enfrentar entonces la afirmación de que pareciera que podemos imaginarnos que fuera de otra manera? Pareciera que podemos imaginar una situación contrafáctica, un ‘mundo posible’ en el que la identidad es falsa.

Kripke subraya que si lo que se quiere decir es que se pueden imaginar circunstancias *que se describirían* como circunstancias en las que *a* no habría sido *b*, lo que debemos preguntarnos es si estas circunstancias imaginadas son circunstancias en las que *a* no habría sido *b* (es decir, si es una situación contrafáctica *correctamente descrita* como una situación en la que *a* no habría sido *b*).

Pero lo que encuentra es que, por ejemplo, la situación en la que resulta que Phosphorus era Marte, y Hesperus era Venus, sería una situación en la que la gente *de ese mundo* habría dicho con verdad ‘*a* no es *b*’; pero debemos describir las cosas en *nuestro lenguaje*, no en el de ellos. Usando las palabras en como las usamos nosotros, lo que sucedería entonces es que, bajo esas circunstancias, Phosphorus en la tarde no estaría en la posición en donde lo veíamos, o alternativamente, que Hesperus en la tarde no estaría en la posición en la cual lo veíamos. Incluso podríamos decir que bajo esas circunstancias, no habríamos llamado “Hesperus” a Hesperus, porque Hesperus habría estado en una posición diferente. Pero aún eso no haría a Phosphorus diferente de Hesperus.⁶ Por su parte, las circunstancias en las que se supone que “*el calor es el*

⁶ Es a propósito de este ejemplo de Phosphorus y Hesperus que Kripke llama la atención sobre el hecho de que algunas veces fijamos la referencia de un término mediante una descripción. Sin embargo, una vez que tenemos fijada la referencia, usamos el término rígidamente para designar, no a cualquiera que hubiera cumplido la descripción, sino a la cosa nombrada (incluso en situaciones contrafáticas en las que la cosa no hubiera cumplido la descripción). El nombre sería usado para nombrar a un objeto que hemos seleccionado por el hecho contingente de que

movimiento de las moléculas” habría sido falso, son circunstancias en las que se da una confusión entre la sensación de calor y el calor.

De esta manera, las circunstancias en las que se piensa que un enunciado de identidad a posteriori habría sido falso no son realmente tales. Lo que nos produce la ilusión de contingencia es el hecho de que hemos identificado a *a* por un hecho contingente, por una propiedad contingente (aparecer en tal parte de cielo a tal hora del día, producir tal o cual sensación en mí). De tal manera que en estos casos sólo tenemos la *ilusión* de contingencia, la ilusión de que el enunciado podría no ser verdadero en algún mundo posible.

Ahora bien, podemos imaginar circunstancias en las que la identidad mente-cuerpo habría sido falsa. Pero el modelo anterior para desvanecer la ilusión de contingencia no sirve en el caso de la identidad mente-cuerpo, pues no se puede decir que seleccionamos/identificamos el dolor por una propiedad contingente. Los términos mentales de estados conscientes y los términos neurofisiológicos son designadores rígidos, seleccionan sus objetos por propiedades esenciales (no por propiedades contingentes o accidentales). Tampoco podemos operar como cuando separamos la sensación de calor del calor, pues no tiene sentido pretender separar lo subjetivo asociado a lo objetivo en el caso de la experiencia de dolor. De manera que el modelo que tenemos para desvanecer la ilusión de contingencia no nos ayuda a concluir en este caso que no nos estamos imaginando la falsedad de esta identidad.

Kripke hace énfasis en que el fisicalista tiene que negar que pudiera haber habido situaciones en las que se hubiera dado el estado consciente sin el estado cerebral. Obviamente, de ser así habría un ‘mundo posible’ o situación contrafáctica en el que el enunciado de identidad psicofísico es falso, y el enunciado no sería necesario (es decir, no sería verdad en todo mundo posible) sino contingente y entonces la identidad sería falsa. El fisicalista tiene que sostener que somos presa de una ilusión al pensar en los

satisface una determinada descripción. Pensar lo contrario sería malinterpretar la relación entre un nombre y la descripción usada para fijar su referencia, considerándolos sinónimos.

zombis, etc. Pero ¿qué modelo podría utilizar el fisicalista si el único que hay a mano no cumple la función requerida?

1.1.1 El Desafío de Kripke

Al final de su ensayo, Kripke deja abierta la pregunta sobre a qué modelo podría recurrir el fisicalista para desvanecer la que sería una ilusión de contingencia de la identidad psicofísica. Esto es, cuál sería la estrategia para descartar las situaciones en las que parece que podemos imaginar que los estados mentales no son idénticos a estados físicos, y probar que se trata de situaciones contrafácticas erróneamente descritas. Pero esta pregunta abierta tiene un dejo levemente retórico, pues Kripke pareciera sugerir que no existe un modelo tal que esté disponible para el fisicalista (cf. Kripke 1972, n. 18), lo cual significaría que cualquier identidad psicofísica sería inevitablemente falsa.

Con todo, ¿qué argumentos se han dado para pensar que el modelo según el cual se identifica algo por medio de una propiedad contingente asociada, es el único modelo para desvanecer la ilusión de contingencia? Lo único que ha ofrecido Kripke es afirmaciones en el sentido de que no se le ocurre qué modelo podría hacer esto, e incluso ha manifestado sus sospechas de que no haya uno; pero esto, por supuesto, no equivale a un argumento.

Y bien ¿qué es lo que debe hacer un modelo tal? Parece necesario decirlo de nuevo, pero esta vez de una forma más explícita: debe permitirnos concluir que algo que pensamos que podemos imaginar (que concebimos que puede ser posible), no lo es (no es realmente algo que podemos imaginar).

No obstante, como veremos con mayor detalle más adelante, algunos de los intentos de responder a este desafío de Kripke operan con una disociación entre lo epistémico y lo ontológico, según la cual el que nos imaginemos una cierta situación no hace que sea posible. Para Joseph Levine (1983, 1993), por ejemplo, la ilusión de contingencia de la identidad psicofísica no equivaldría al hecho de que creamos que podemos imaginar una situación contrafáctica que de hecho no podemos imaginar. Para él *sí* podemos imaginar tal situación, sólo que su imaginabilidad o concebibilidad no significan que pueda ser posible; es decir, la ilusión de contingencia no se derivaría del hecho de que *en efecto*

me puedo imaginar tal situación contrafáctica, sino a que creo *erróneamente* que dado que me la puedo imaginar, puede haber un mundo posible en el que la identidad psicofísica es falsa.⁷

Según lo anterior, pareciera que podemos trazar dos caminos diferentes para abordar el asunto que le plantea Kripke al fiscalista. Uno de ellos aseveraría que *no es concebible que un enunciado de identidad verdadero pudiera ser falso*, mientras que el otro se comprometería con lo contrario, es decir, con que *se puede concebir que un enunciado de identidad verdadero sea falso* aunque rechazando obviamente la consecuencia de que esto involucre la contingencia (ergo, falsedad) del enunciado de identidad en cuestión. Levine, como veremos, recorre el segundo camino; a Kripke lo encontramos ciertamente en el primero.

Para Kripke, la diferencia que hay entre la concebibilidad o no concebibilidad de una situación es central: si realmente la hay, entonces existe un 'mundo posible' en el que la identidad en cuestión es falsa; pero si una identidad es verdadera, no puede ser falsa en ningún mundo posible (recordemos que 'necesario' es intercambiable por 'verdadero en todos los mundos posibles'). Si la identidad es verdadera-necesaria, no podemos entonces concebir ese mundo posible.

Para Kripke, *debe* haber un momento del modelo que disipa la ilusión de contingencia en el cual, después del análisis de algún supuesto contraejemplo al enunciado de identidad, se llegue a una conclusión como "esto no iría en contra de decir que en esta situación contrafáctica *a* seguiría siendo *b*" o "*x* no cuenta, entonces, como un caso en el que me puedo imaginar que *a* no es *b*". El asunto para Kripke radica en que disipar la ilusión de contingencia es probar que realmente no se trata de circunstancias en las que *a* no habría sido *b*. En general, debe concluirse que se trata de confusiones, de posibilidades mal descritas.

⁷ En este sentido, mientras que para Kripke la 'ilusión de contingencia' de una identidad equivale a la falsa creencia de que se puede concebir que la identidad fuera falsa, para Levine, dicha ilusión sería equivalente a la falsa creencia de que en la medida que se puede concebir que la identidad sea falsa, entonces es falsa.

Sin embargo, en las afirmaciones de Kripke muchas son las cosas que quedan sin aclarar. Decir que algo que me imagino es una ‘posibilidad mal descrita’ nos deja en la oscuridad acerca de cuestiones significativas. Con ‘posibilidad mal descrita’, Kripke parece estar diciendo que lo que creo que me puedo imaginar no es realmente lo que me estoy imaginando; pero esto debería sorprendernos, pues comúnmente creemos que quien está en una mejor posición para determinar cuándo alguien se está imaginando algo o no, es quien se lo imagina. De lo contrario, no queda claro qué o quién es exactamente el que determina qué es aquello que me estoy imaginando, o cuándo he logrado imaginarme algo y no simplemente he creído erróneamente haberlo imaginado. Esto requiere decantar mucho más cuestiones bastante complejas, como lo que significa concebir o imaginarse algo: no se nos dice nada sobre si para concebir una situación contrafáctica basta con hacerse alguna especie de imagen mental o si basta con formularlo lingüísticamente, si esa formulación lingüística o imagen mental puede ser vaga o tiene que ser precisa, etc.

Por su parte, Levine pareciera no encontrar nada significativo en el hecho de que una situación se pueda imaginar, al menos en lo que respecta a las posibilidades que pueden darse o no en el mundo real. Aproximaciones como la de Levine ponen en tela de juicio la intuición de Kripke sobre el poder de nuestras intuiciones, las cuales pueden no ser un indicador fiable de lo que es posible. La crítica parte de cuestionar la idea de que las construcciones que pueda hacer nuestra mente –que dependen de nuestras capacidades cognitivas y epistémicas– sean un indicador fiable de las posibilidades que pueden o no darse en la realidad objetiva independiente de nosotros. ¿Cuál es el *argumento* para pensar que todas las situaciones que podrían presentarse en el universo felizmente se pueden calcular por operaciones de nuestro pensamiento?

Vale la pena destacar aquí una serie de elementos clave que aparecen en la argumentación de Kripke. En primer lugar, si nos fijamos en los rasgos comunes a los casos en los cuales se presenta la ‘ilusión de contingencia’ (“*Hesperus=Phosphorus*”, “*Calor=Movimiento de Partículas*”), podemos ver que en todos su verdad sólo se puede conocer *a posteriori* debido a su carácter empírico. En todos los casos de enunciados de identidad cuya verdad sólo puede conocerse *a posteriori*, pareciera que podemos

imaginarnos una situación contrafáctica o ‘mundo posible’ en que hubiera sido de otra manera⁸. Este rasgo común a los enunciados de identidad empíricos podría dar cuenta de por qué pareciera que podemos formular situaciones contrafáticas que los falsean: se relacionaría más con nuestra situación epistémica (con el hecho de que sólo podemos conocer su verdad *a posteriori*) que con la naturaleza misma de la realidad.

Lo anterior le podría proporcionar a Levine una línea argumentativa para explicar por qué nos imaginamos ciertas situaciones aunque los enunciados de identidad relacionados puedan seguir siendo verdaderos y necesarios.⁹ Si existiera una relación entre concebibilidad y aposterioridad, o, más aún, si lográramos ligar la concebibilidad con la aposterioridad, y tenemos en cuenta la afirmación de Kripke de que *el hecho de que un enunciado sea a posteriori no implica que no pueda ser necesario*, podríamos llegar a la conclusión de que *del hecho de que se pueda concebir la falsedad del enunciado no se sigue que no sea necesario*.

Por otro lado, recordemos también que el mismo Kripke subraya la diferencia entre el ámbito ontológico y el epistémico, en especial cuando hace sus observaciones hacia Quine y Marcus (véase la nota 13 de *“Identidad y Necesidad”*, Kripke 1972). A propósito de las causas de la ilusión de contingencia en identidades como “calor=movimiento de las partículas”, “Hesperus=Phosphorus”, “Marco Tulio=Cicerón”, asegura: “*se relacionan con la confusión de la cuestión metafísica de la necesidad (...) con la cuestión epistemológica de su aprioridad*” (1972, n.14). Levine pareciera simplemente estar anotando lo mismo a propósito de la identidad psicofísica. ¿Cuál es el argumento por el cual no deberíamos pensar que es esto lo que sucede con la identidad psicofísica?

Si seguimos la citada indicación kripkeana de distinguir lo epistemológico de lo ontológico (de separar la cuestión de la necesidad de la cuestión de la aprioridad/aposterioridad), el que una identidad sea contingente y falsa (cuestión metafísica) no puede depender meramente del hecho de que me pueda imaginar una circunstancia en la que no se da (cuestión epistemológica).

⁸ Otra confusión que opera en estos casos, que como veremos resulta bastante relevante para Ned Block y Robert Stalnaker (1998), es la que hay entre lo que *nosotros* diríamos en una situación contrafáctica y cómo la describiría la gente en esa situación.

⁹ Evidentemente, esta línea argumentativa depende en gran parte de que se haya hecho una aclaración satisfactoria de las cuestiones planteadas atrás sobre la noción de concebibilidad.

Pero según dijimos, la importancia que tiene para Kripke el asunto de la concebibilidad (i.e., que se pueda concebir una circunstancia en la que el enunciado de identidad no se da) precisamente es que si realmente la hay, entonces existe un 'mundo posible' en el que un enunciado es falso, por lo que sería contingente y no necesario. Si la identidad es necesaria, no deberíamos poder concebir ese mundo posible. Así pues, parece que Kripke estaría en riesgo de caer en este mismo error de confundir los asuntos de la ontología con los de la epistemología.

Pero ¿cuál es la razón por la que Kripke resulta pensando que a cada situación que se puede concebir le corresponde un 'mundo posible'? De acuerdo con Levine, esto sucede porque cree que nuestras intuiciones sobre lo que es concebible tienen cierto poder sobre lo que es posible, porque cree que la concebibilidad es una guía segura para la posibilidad. Pero esto deja a Kripke en el siguiente dilema: o bien se compromete con la idea de que la concebibilidad es una guía para la posibilidad, o bien se compromete con la idea de que hay que diferenciar las cuestiones epistemológicas de las ontológicas. En la medida en que no se ha dado ningún argumento a favor de lo primero, podríamos conservar nuestro compromiso con lo segundo.

Volviendo al desafío de Kripke, el asunto es si un modelo alternativo para explicar la 'ilusión de contingencia' de un enunciado de identidad psicofísica podría aceptar que se puede concebir su falsedad. Por ahora lo que podemos decir es que si lo rechazáramos sería sobre bases que no son incuestionables (básicamente, la idea de que la concebibilidad es una guía segura para la necesidad). Pero si seguimos la invitación de distinguir lo epistemológico de lo ontológico, no parece haber ninguna razón para rechazar la concebibilidad (pues no estaríamos obligados a aceptar la idea de que la concebibilidad es una guía para la posibilidad, ni ninguna idea semejante en la que le atribuyamos a nuestras intuiciones poder alguno sobre las posibilidades que existen en el mundo real). Resulta razonable que lo que yo considero una situación contrafáctica posible no lo sea.

Pero ¿realmente una postura tipo Levine constituye el santo grial para el fisicalista? Hace un momento dije que, para Levine, no sería muy significativo el hecho de que una situación se pueda imaginar; esto no es del todo cierto o preciso. Podemos adelantarnos

un poco y advertir una consecuencia que la posición de Levine acarrea: de la concebibilidad de ciertas situaciones que falsean la identidad psicofísica se sigue la existencia de una brecha explicativa insalvable, sin ninguna superación posible ¿no sería entonces el de Levine un flaco favor al fisicalismo?

Quiero finalizar este apartado anotando que la de Kripke es una de las principales líneas de argumentación contra de la teoría de la identidad y el fisicalismo acerca de lo mental, que se ha conocido como el '*argumento modal*', o las '*objeciones de concebibilidad*'. Desde luego, la primera aparición de este tipo de objeciones es bastante anterior a Kripke, como él mismo reconoce, y ya se encuentra formulado en las meditaciones metafísicas de Descartes y es ampliamente discutido en las objeciones y respuestas a las *Meditaciones*.

1.2 NAGEL Y LA SUBJETIVIDAD DE LA EXPERIENCIA

Luego de haber explorado una de las principales y más fuertes argumentaciones contra el fisicalismo, es oportuno examinar otra de las reflexiones más difundidas sobre la relación entre las experiencias conscientes y cualesquiera estados físicos. Si bien dichas reflexiones no están dirigidas a mostrar la falsedad del fisicalismo, sí constituyen una problematización definitiva del alcance de la tesis fisicalista. Se trata de la argumentación presentada por Thomas Nagel en su artículo de 1974 "*What it is like to be a bat*" (Nagel, 1974) en el que pone en evidencia una dificultad significativa que se le presentaría al teórico de la identidad, aunque en una dirección muy diferente a la de Kripke.

Para Nagel, ninguno de los ejemplos reduccionistas usuales nos ayuda a entender verdaderamente la relación entre la mente y el cuerpo, por lo cual realmente no tenemos ninguna idea de cuál podría ser la explicación de la naturaleza física de un fenómeno mental. De acuerdo con Nagel, todas las aproximaciones al tema de la conciencia la han concebido sistemáticamente de manera errónea, en especial porque ignoran lo que hace único el problema mente-cuerpo. Una vez se dimensiona correctamente aquello que hace único el problema mente cuerpo, resulta evidente –continúa Nagel– que ninguna de las concepciones conocidas de reducción es aplicable a la conciencia.

Según Nagel, que un organismo tenga experiencia consciente, significa fundamentalmente que *hay algo que es ser como ese organismo*. Nagel observa de inmediato que este es el carácter subjetivo de la experiencia, y que este carácter que no es capturado por ninguno de los análisis reductivos de lo mental en la medida que todos ellos son lógicamente compatibles con su ausencia. Los análisis causales y funcionales, por principio, dejan algo por fuera; Nagel aclara que no pretende decir que las experiencias conscientes no tengan poder causal o un rol funcional, sino que esto no agota su análisis, no dice nada sobre cómo es tener tal o cual experiencia consciente.

Enseguida, pasa a desarrollar con más detenimiento su argumento a partir de un ejemplo. Si yo tratara de imaginar cómo es ser un murciélago estaría restringido a los recursos de mi propia mente, y esos recursos parecen inadecuados para la tarea pues mis experiencias no son como las del murciélago. Si pensara que de lo que se trata es que de que mi neurología se metamorfosee en la del murciélago, nada me permite siquiera imaginar cómo serían las experiencias de una etapa de dicha metamorfosis. En suma, no podemos formar más que una concepción sumamente esquemática de cómo es ser un murciélago: simplemente podemos pensar que sus experiencias tienen un carácter subjetivo específico que está más allá de nuestra habilidad para concebir.

Nagel afirma que el punto central que quiere subrayar con este ejemplo es el de la divergencia entre los tipos de concepción subjetivo y objetivo: independientemente del status ontológico que tengan los hechos acerca de 'como qué es ser un x' estos hechos *incorporan un punto de vista particular*.¹⁰ Ahora, si los hechos de la experiencia son accesibles sólo desde un punto de vista, entonces ¿cómo puede ser revelado el auténtico carácter de las experiencias en la operación física de ese organismo? Generalmente –señala Nagel– una reducción se define como un movimiento hacia una visión objetiva, más exacta, de la real naturaleza de las cosas, que se logra reduciendo la dependencia a puntos de vista sobre el objeto de investigación. Pero eso no se puede hacer con la experiencia consciente, pues la idea de moverse de la apariencia a la realidad no tiene sentido aquí. No podemos acercarnos a la naturaleza real de la experiencia humana dejando atrás el punto de vista que de hecho constituye su esencia.

¹⁰ Inmediatamente Nagel hace la claridad de que no se trata de la "privacidad" de la experiencia: el punto de vista en cuestión no es una 'instancia' sino un 'tipo'.

No cabe duda, concede Nagel, de que existen conceptos que capturan la naturaleza objetiva intrínseca de algo. En este sentido, es posible una reducción de algunas cosas (las cuales no requieren puntos de vista), pero no es posible la reducción de la experiencia (pues, sin duda, esta requiere un punto de vista). Podemos pensar en el carácter objetivo de la luz ¿pero qué sería el carácter objetivo de la experiencia en cuanto tal? Si se elimina lo subjetivo ¿qué queda de la experiencia?

Nagel señala que si ha de defenderse el fisicalismo, se le deben dar a los rasgos fenomenológicos mismos una explicación física. Pero recalca el punto hecho igualmente por Kripke contra el fisicalismo, que modelos ‘des-subjetivizantes’ como el del *calor/movimiento de las partículas* no tiene ninguna aplicación en este caso.

Ahora bien, Nagel observa que si los procesos mentales son procesos físicos, entonces hay algo que es como estar en ciertos procesos físicos. Pero el mismo Nagel admite que el hecho de que no entendamos cómo es posible que haya algo que es estar en tal o cual estado físico, no significa necesariamente que el fisicalismo sea falso. Esta incapacidad explicativa no prueba nada acerca del fisicalismo: lo más que puede decirse sobre el fisicalismo es que se trata de una posición que no podemos entender porque no tenemos una concepción de cómo podría ser verdad.¹¹

Un antifisicalista podría replicar que lo que dice el fisicalismo es bastante claro y que si no podemos formarnos una concepción de él debe ser porque es falso. Pero Nagel enfatiza que no es nada claro. Supongamos que $X=Y$, sabemos cómo refieren tanto “X” como “Y” –conocemos las cosas a las que refieren– y tenemos además una vaga idea de cómo pueden converger dos senderos referenciales en una sola cosa. Pero cuando los dos términos son muy disímiles –como en el caso de la identidad psicofísica– puede no ser tan claro cómo podría ser cierto, podríamos no tener ni una vaga idea de cómo

¹¹ La carencia de semejante concepción no es suficiente para negar la posible verdad de la afirmación. Recordemos que Nagel ha afirmado que cree en la existencia de hechos más allá del alcance de los conceptos humanos. Se trata de una especie de limitación humana de principio, muy a semejanza de lo que plantea Collin McGinn (McGinn 1991). Nagel cree que hay hechos que no podrían ser representados o comprendidos *nunca* por los seres humanos, hechos que no consisten en la verdad de proposiciones expresables en un lenguaje humano.

podrían converger los dos senderos referenciales de expresiones tan disímiles, o en qué clase de cosas podrían converger.

Nagel recalca que en este momento no tenemos ningún esbozo de una concepción de cómo el fisicalismo podría ser cierto, es decir, no tenemos idea de cómo un evento mental y uno físico podrían referirse a la misma cosa (y las analogías con identidades teóricas no nos ayudan a obtenerla, pues hacen caso omiso de lo subjetivo, entre otras). Pero admite que aunque parezca raro, tenemos evidencia de que algo que realmente no podemos entender es verdad: retoma aquí el planteamiento de Davidson según el cual si los eventos mentales tiene causas y efectos físicos deben tener también descripciones físicas, incluso aunque no tengamos –y, de hecho, aunque no podamos tener– una teoría psicofísica general; recuerda también que la neurofisiología nos ha dado razones para creer, por ejemplo, que las sensaciones son procesos físicos, aunque no estemos en posición de entenderlo.

Si resultara que el fisicalismo es verdad, Nagel mismo piensa que es posible construir una explicación de la ‘ilusión de contingencia’ (planteada por Kripke y discutida anteriormente). Esta explicación se basa de la distinción entre imaginación ‘simpática’ e imaginación ‘perceptual’. Cuando tratamos de imaginar un estado mental ocurriendo sin su estado cerebral asociado hacemos dos cosas: primero, imaginamos ‘simpáticamente’ la ocurrencia del estado mental (es decir, nos ponemos a nosotros mismos en un estado que se le parece mentalmente). Al mismo tiempo, intentamos imaginar perceptualmente la no-ocurrencia del estado físico asociado (es decir, nos ponemos a nosotros mismos en otro estado –que no tiene ninguna relación con el anterior– que se parece al estado en que estaríamos si percibiéramos la no-ocurrencia del estado físico). En la medida en que la imaginación de los rasgos físicos es perceptual y la imaginación de los rasgos mentales es simpática, parece que podemos imaginar cualquier experiencia ocurriendo sin su estado cerebral asociado, y viceversa. La relación entre ambos parecerá contingente incluso si es necesaria, a causa de la independencia de los disímiles tipos de imaginación.

1.2.1 Subjetividad y Fisicalismo

Examinemos de nuevo el sentido exacto de la pregunta “¿cómo es ser un murciélago?”. Nagel insiste en que el asunto es que no podemos hacernos una idea de cómo es esa forma de experiencia perceptual, que no podemos hacernos una idea del punto de vista que tiene un murciélago como sujeto de esos estados, y que para esto no tiene utilidad alguna hacer una extrapolación de cómo es la vida interna del murciélago a partir de nuestro propio caso. Esto es lo que significa que no podemos saber cómo es ser un murciélago.

Más adelante (Nagel, 1974 n. 6), Nagel apunta que una forma más precisa de entender la pregunta “cómo es ser un murciélago” no es ni siquiera como “a qué (*en nuestra experiencia*) se parece” sino “cómo es para el sujeto mismo”. Creo, sin embargo, que la precisión que se pretende hacer en este pie de página acentúa una ambigüedad que sólo es superada más adelante en el texto. Como indica el mismo Nagel, cuando hablamos del punto de vista del sujeto, podemos estar hablando una instancia (es decir, el punto de vista de un individuo particular, accesible sólo a ese individuo) o de de un tipo (el punto de vista de cierta clase de individuos, accesible a los individuos pertenecientes a dicha clase). Correspondientemente, existen dos interpretaciones posibles, significativamente distintas de la pregunta *¿cómo es ser un murciélago?* si lo que estoy preguntando es *¿cómo es para el sujeto mismo (es decir, el murciélago mismo) ser un murciélago?*

Por un lado, podría leerse como la observación de que no puedo tener el punto-de-vista-instancia del murciélago (Kathleen Wider (Wider, 1990) parece hacer esta interpretación); por otro lado, podría interpretarse como la observación de que no puedo tener su punto-de-vista-tipo. En el primer caso, no solamente no valdría la pena la aclaración en contra de la privacidad que se hace más adelante, sino que se trataría de un punto menos interesante para el problema mente-cuerpo, uno que ya han hecho en el pasado muchos filósofos de la mente: que no puedo saber cómo es ser tal o cual individuo particular, porque no soy ese individuo (de hecho, ni si quiera puedo saber si hay algo que es ser ese individuo); este es un punto más cercano al problema de otras mentes que al punto que Nagel pareciera estar tratando sobre la subjetividad y la objetividad. Es más, algunos dirían que se estaría haciendo una observación obvia, un truísmo: no puedo tener el punto de vista del murciélago, porque no soy el murciélago.

Posteriormente Nagel aclara que se refiere al segundo caso, notando que la mencionada ambigüedad podría desorientarnos y llevarnos a pensar que su argumentación apunta al problema de otras mentes y no al asunto de la existencia de una barrera epistemológica para imaginar cierta clase [tipo, no instancia] de experiencias conscientes. Hace un momento mencionábamos la afirmación de Nagel con respecto a que el núcleo del ejemplo del murciélago es la divergencia entre los tipos de concepción subjetivo y objetivo, y en particular la observación de que los hechos acerca de ‘cómo es ser un x ’ incorporan un punto de vista particular. El punto de Nagel es que para entender “cómo es ser x ” debe tomarse el punto de vista [tipo] de x ; si no se lo puede tomar, no se puede saber cómo es ser x . Por ende, se requiere algo eminentemente subjetivo (no objetivo) para saber “cómo es ser x ”.

Una vez precisada la pregunta, podemos afirmar que el principal aporte que hacen estas reflexiones de Nagel sobre la relación entre las experiencias conscientes y cualesquiera estados físicos es la observación de que el carácter subjetivo de la experiencia no es capturado por ninguno de los análisis reductivos de lo mental. La prueba irrefutable de esto es que todos estos análisis reductivos son lógicamente compatibles con la ausencia de subjetividad; esto significa que análisis de lo mental como el fisicalista o funcionalista son conceptualmente independientes de un análisis del carácter subjetivo de lo mental (¿cómo podrían, pues, capturarlo si no lo implican de ninguna manera?).

Podría decirse incluso que por principio estos análisis reductivos suprimen el carácter subjetivo, pues tienen la pretensión de eliminar lo que le es esencial: el punto de vista del sujeto (en otras palabras, podría decirse que un análisis reductivo –al menos un análisis reductivo de tipo funcionalista o fisicalista– fracasa o dejar de ser reductivo si no elimina todos los rasgos subjetivos). Más aún, resulta imposible entonces reducir la experiencia consciente (por ejemplo, a la terminología de la ciencia física), pues el proceso de objetivización implicado por la reducción es incompatible con la comprensión de la experiencia consciente misma: no quedaría nada de lo reducido. Todo fenómeno subjetivo está esencialmente conectado con un punto de vista particular, y parece inevitable que una teoría física objetiva abandone ese punto de vista; si la reducibilidad se ve como equivalente a la objetivización, se sigue que lo que es subjetivo es irreducible (no objetivizable).

En relación con este punto, recordemos que en cierto momento de la argumentación Nagel señala que para hacer del fisicalismo una postura defendible “*se le debe dar a los rasgos fenomenológicos mismos una explicación física*” (Nagel, 1974 p.437). Al observar detenidamente, no es completamente claro lo que esto quiere decir: podría interpretarse como (i) la exigencia de que los rasgos fenoménicos sean descriptible en el vocabulario de la física; o podría expresar (ii) la necesidad de que los rasgos fenoménicos deban poderse incluir como entidades del dominio de lo físico, esto es, concebidos como entidades físicas.

Si Nagel estuviera pensando algo de la primera especie, estaría diciendo –palabras más, palabras menos– que el fisicalismo es falso, pues vimos que es enfático en negar que los rasgos fenoménicos puedan ser explicados en el vocabulario objetivista de la física. Pero él mismo señala que esta incapacidad explicativa del fisicalismo no acarrea la negación del fisicalismo, por lo cual en el pasaje citado parecería estar pensando en algo de la segunda especie. Esta interpretación se vería reforzada por el hecho de que Nagel se refiera al argumento de Davidson y el argumento causal como ‘evidencias’ de que los rasgos fenoménicos podrían ser incluidos como entidades del dominio de lo físico, pese a que no podamos entenderlo por completo.¹²

¿Debería esto último permitir al fisicalista respirar tranquilo ante las consideraciones de Nagel? De ningún modo. Como mostrará la argumentación de Levine, y cómo ya ha venido anticipando la argumentación de Nagel, se siguen importantes consecuencias de la irreductibilidad del carácter subjetivo de la experiencia, en especial la existencia de una brecha entre lo que Nagel ha llamado “dos tipos de concepción”.

En cierto momento, Nagel se cuestiona con un aire un poco retórico “*si los hechos de la experiencia (...) son accesibles sólo desde un punto de vista, entonces es un misterio cómo puede ser revelado el verdadero carácter de las experiencias en la operación física*”

¹² Otro indicio de que Nagel no está formulando una negación del fisicalismo podría ser el hecho de que desde el comienzo del ejemplo del murciélago parta de que su diferencia experiencial está relacionada con una diferencia en su neurofisiología (con la manera en que “sus cerebros están diseñados”).

de ese organismo" (Nagel, 1974 p.442). El dejo retórico radica en que ya hemos visto que la respuesta a este interrogante es que no puede hacerlo.

No obstante, y contrario a como se ha leído sistemáticamente a Nagel, pienso que su argumentación ofrece algunas pistas que resultan claves para el fiscalista. En primer lugar, Nagel ha anotado que cuando dos términos son muy disímiles puede que no veamos cómo podrían converger sus dos senderos referenciales en la misma cosa, o en qué clase de cosas podrían converger. Algo como esto podría ser lo que da cuenta de nuestra incapacidad de entender que los términos mentales y algunos términos físicos se refieran a lo mismo.

Desde Descartes, se ha pensado que mientras que cualquier objeto físico es, en principio, directamente observable por todos, sólo quien es sujeto de determinadas experiencias conscientes puede acceder a ellas; este tipo de cosas le han sugerido a muchos que las experiencias no tienen nada en común con los objetos físicos. Pero un contraste como este sólo se puede sostener en una visión anticuada de lo físico: los campos electromagnéticos o las partículas subatómicas elementales no poseen ninguna de las propiedades que nuestros sentidos revelan directamente (i.e., no son 'directamente observables por todos'). En esta medida, parece que la semejanza entre aquello a lo que los términos mentales y los términos físicos refieren puede deberse a una concepción ingenua de lo físico, es decir, a una noción anticuada de lo que es una entidad física.

Además, de hecho tenemos una idea –aunque sea absolutamente esquemática– de cómo los dos senderos referenciales podrían converger, proporcionada por la evidencia que el mismo Nagel cita (el argumento causal y el argumento de Davidson). Esta idea esquemática bien podría contar como el trasfondo conceptual o teórico –si bien mínimo, rudimentario– requerido por Nagel. El resto de la explicación podría llegar después, o incluso no llegar nunca; lo que cuenta es que no tengamos un problema de principio para pensar que es el caso que los dos senderos referenciales convergen –de nuevo, aunque no lo entendamos: no tenemos que entenderlo de una manera absolutamente precisa para permitirnos concluir que es el caso–.

En segundo lugar, y de manera conexa con esto último, otra de las reflexiones que encontramos en el artículo de Nagel es que hay *hechos* que no consisten en la verdad de proposiciones expresables en un lenguaje humano. Aunque no podamos explicar un hecho, aunque no entendamos cómo puede ser verdad, no significa que no sea verdad; más aún, podemos tener *evidencia* de algo que no podemos entender.

Un rasgo importante que vale la pena tener en cuenta en esta idea nageliana, es que guarda continuidad con la invitación kripkeana de diferenciar las cuestiones epistémicas de las metafísicas. Para verlo con más claridad, digámoslo una vez más, añadiendo un par de énfasis: aunque no *entendamos* cómo algo puede ser verdad, no significa que no sea verdad. Algo que no entendemos puede ser el caso; nuestra incapacidad (temporal o permanente) para comprenderlo no afecta su verdad.

Por todo lo anterior, podríamos pensar que sufrir ciertos procesos físicos se siente de cierta manera (usando la terminología de Nagel, que ‘hay algo que es sufrir ciertos procesos físicos’), aunque eventualmente pueda resultar extraño y misterioso. Así, cuando pregunta “¿*tiene sentido decir que las experiencias tienen un carácter objetivo?*” (p. 448), podemos responder algo en del siguiente talante: si los estados mentales son estados físicos (digamos, si la identidad “dolor=activación de las fibras C” fuera verdad) entonces en un sentido las experiencias tendrían un carácter objetivo: el carácter objetivo de la activación de las fibras C.¹³

Además, si Nagel está siguiendo la exhortación de distinguir las cuestiones de epistemología de las cuestiones de ontología, podemos suponer que admitiría que la brecha que denuncia entre los *modos de concepción* subjetivo y objetivo, es una brecha que se da entre maneras de concebir – *por tanto*, en el ámbito epistémico– y no entre dominios ontológicos –ámbito metafísico–.

¹³ El papel que en este planteamiento juegan lo que Nagel llama “dos modos de concepción” y, en general, la idea de que podemos tener más de una forma de acceso epistémico a una cosa, se abordarán con mucho más detalle en lo que sigue.

1.3 JACKSON Y LA “*INFORMACIÓN FÍSICA*”

Otro de los argumentos más discutidos contra el fisicalismo acerca de la conciencia fenoménica es el célebre Argumento del Conocimiento de Frank Jackson (Jackson, 1982), en el que se pretende establecer que la experiencia consciente involucra propiedades no-físicas. El argumento se basa en la idea de que alguien que tiene toda la información física acerca de la experiencia de colores podría no saber cómo es tener una experiencia de color.

De acuerdo con Jackson, la información proporcionada por la ciencia natural (información “*acerca del mundo, acerca de los procesos que suceden en el cerebro y de la manera en que se relacionan con el entorno*” p. 127) es lo que denomina ‘*Información Física*’, por la cual se define lo que es una propiedad física o un proceso físico. En este sentido, afirma, podemos definir el fisicalismo como la tesis de que toda la información (correcta) es información física.

Sin embargo, plantea que no existe cantidad alguna de información física suficiente para incluir ciertos rasgos de las experiencias conscientes. Nada que se pueda estructurar en términos de información física captura las experiencias conscientes; esto, desde luego, significa que el fisicalismo es falso.

Las premisas del anterior razonamiento son, según Jackson, verdades intuitivamente obvias; de hecho comenta que toda la argumentación que va a presentar contra el fisicalismo contiene premisas de este tipo. Además, se anticipa a señalar que el argumento causal al que de una manera u otra siempre acude el fisicalista en busca de refugio, puede ser demolido mediante un recurso al epifenomenalismo.

El Argumento del Conocimiento, como lo llama Jackson, se apoya en dos experimentos mentales. El primer caso que presenta es el de Fred, una persona que distingue colores que nosotros no podemos distinguir. En este escenario, la investigación de las bases fisiológicas de su habilidad revela que su sistema óptico es capaz de separar dos grupos de longitud de onda que nosotros no podemos separar. Pero no existe cantidad suficiente de información física acerca del cerebro y sistema óptico de Fred que nos permita saber cómo es la experiencia que tiene cuando ve los dos colores diferentes. Se sigue que hay algo que no sabemos acerca de las experiencias de Fred aunque tengamos toda la

información física acerca de él. Y si no puede ser que supiéramos todo acerca de Fred, aunque supiéramos la información física acerca de él, entonces el esquema fisicalista deja algo por fuera.

El segundo caso presentado por Jackson es el famoso caso de Mary, que podríamos decir que es un caso inverso al anterior, pues en este caso somos nosotros los que tenemos experiencias de una clase que es desconocida para ella.

Para Jackson es obvio que Mary aprenderá algo cuando salga de su habitación en blanco y negro. Lo cual hace inevitable que su conocimiento anterior (todo el conocimiento físico) sea incompleto, ergo hay más que información que información física.

La conclusión es que en ambos casos algunos rasgos de las experiencias conscientes quedan por fuera de la teoría fisicalista: se puede tener toda la información física sin tener toda la información que hay que tener sobre las experiencias conscientes.

Jackson extrae incluso un corolario modal de la argumentación que acaba de presentar: decir que no existe cantidad suficiente de información física acerca de otra persona que implique lógicamente que tiene conciencia significa que hay un mundo posible con organismos físicamente como nosotros pero que no tienen conciencia. Estos organismos son copias físicas, pero carecen de algo, que por ende no es físico.

Pese a que deriva corolarios modales del argumento que acaba de presentar, Jackson subraya que se trata de un argumento diferente del dado por Kripke y diferente también del argumento de Nagel.

En relación con el argumento de Nagel, Jackson enfatiza que el punto que quiere señalar con el Argumento del Conocimiento no es que no sepamos cómo es *ser* Fred, sino que ignoramos una propiedad acerca de su experiencia (y que podríamos conocer esa propiedad sin saber cómo es *ser* Fred). Observa que no existe cantidad suficiente de información de ninguna clase que equivalga a un 'conocimiento de Fred desde adentro'; mientras que la cualidad especial de la experiencia que ignoramos es un hecho acerca de ella, hecho que el fisicalismo deja por fuera porque no existe cantidad suficiente de información física que lo capture.

La introducción del tema del epifenomenalismo parece una anticipación a la réplica del fisicalista en la que se aduce que si el fisicalismo es falso tenemos serias dificultades para explicar el asunto de la interacción psicofísica. De este modo, Jackson empieza matizando su postura al decir que no concebirá el epifenomenalismo como la afirmación de que los *estados* mentales son ineficientes causalmente con respecto al mundo físico, ni como la afirmación de que lo mental es totalmente ineficiente causalmente. Su postura consistirá más bien en la tesis de que la posesión o carencia de ciertas *propiedades* de ciertos estados mentales (los *qualia*), no hace ninguna diferencia en el mundo físico (un *qualia* puede hacer una diferencia en otros estados mentales, aunque no en nada físico).

Paso seguido, discute tres razones por las que usualmente pensamos que los *qualia* son causalmente eficientes sobre lo físico:

1. Es común pensar que el carácter doloroso del dolor (A) es en parte responsable de que el sujeto busque evitar el dolor (B). Sin embargo, cualquier hipótesis de la forma 'A causa B' puede ser derrumbada por una teoría que muestre que los dos son distintos efectos de un proceso causal común subyacente. El epifenomenalista puede decir que la conexión entre el carácter doloroso y la conducta es que simplemente ciertos sucesos en el cerebro los causan a ambos.
2. Igualmente se arguye que si todos nuestros rasgos fueron seleccionados durante la evolución, podemos asumir que los *qualia* fueron seleccionados de esta manera y, por tanto, serían conducentes a la supervivencia. Pero difícilmente nos podrían ayudar a sobrevivir si no hicieran nada.

Jackson replica que lo que realmente se sigue de la teoría de Darwin es que cualesquiera características seleccionadas son o bien conducentes a la supervivencia (necesariamente con poder causal), o bien un sub-producto de una propiedad tal (no necesariamente con poder causal): el estado cerebral asociado con el *quale* sería el que evoluciona, mientras que el *quale* mismo sería un subproducto.

La evolución, según Jackson, estaría llena de estos subproductos (algunos de los cuales podrían incluso no ser benéficos). Cita el ejemplo de los osos polares, quienes han desarrollado espesas capas de piel y pelo para mantenerse calientes, incluso aunque esto tiene el efecto colateral perjudicial de que los hace demasiado pesados. La pesadez de la espesa capa de piel y pelo del oso polar hace parte de las propiedades del mismo, y se deriva de una propiedad conducente a la supervivencia,

pero claramente esta es una propiedad que no fue seleccionada para garantizar la supervivencia.

3. Generalmente pensamos que la conducta de una persona proporciona razones para creer que tiene *qualia*, pues pensamos que dicha conducta es causada por los *qualia*. Pero, desde luego, un epifenomenalista no considera la conducta como resultado de los *qualia*, por lo que no puede deducir nada acerca de los *qualia* de un sujeto partir de su conducta.

Jackson replica que, en efecto, el epifenomenalista plantea que los *qualia* son más bien efectos de eventos cerebrales, que no causan nada físico aunque sean causados por lo físico. Pero esto no le impide inferir los *qualia* de otros a partir de su conducta, pues si la conexión entre el carácter doloroso y la conducta es que son causados por ciertos sucesos cerebrales, entonces a partir de la conducta de otros se pueden inferir sus causas cerebrales y a partir de estas se pueden inferir sus *qualia*.

Jackson finaliza señalando que el fisicalismo es un enfoque excesivamente optimista sobre nuestros poderes epistémicos, y plantea que debemos admitir que es muy probable que haya una parte del esquema completo de las cosas que nunca lleguemos a conocer o comprender por más que evolucionemos. Sólo entendemos lo que necesitamos entender para sobrevivir; debe haber asuntos que caen por fuera de nuestra comprensión.

1.3.1 “INFORMACIÓN FÍSICA”

Unas de las primeras y más difundidas críticas que recibió el argumento de Jackson se relaciona con la ambigüedad de la noción de “información física” (v.g., Horgan 1984). Esta noción parece compatible con la siguiente definición:

1. Información Física: aquella que es dada en términos del lenguaje de la ciencia (objetivista, extensional, de tercera persona).

Pero también con la siguiente definición:

2. Información Física: aquella que es acerca de entidades físicas.

Mientras que, por definición, la información física en el primer sentido debe estar expresada en términos propios de ciencias como la física, en el segundo sentido la

información física puede estar dada en otros términos (como los de la economía, o los del lenguaje mentalista).

Así, cuando Jackson señala que la ciencia natural “*ha proporcionado una gran cantidad de información acerca del mundo*” (Jackson 1982, p. 127), parece estarse refiriendo a la información dada en términos de un lenguaje extensional, objetivista, que presenta sus referentes desde el punto de vista de la tercera persona.

Pero la confusión entre los dos sentidos de información física se hace más patente cuando Jackson define el fisicalismo como la afirmación de que “*toda la información (correcta) es información física*” (Jackson 1982, p. 127). Naturalmente, el fisicalista aceptaría que se definiera su posición de esta manera únicamente en el segundo sentido de ‘información física’, no en el primero. De hecho, el fisicalista no está obligado a mantener ningún compromiso con esta definición de fisicalismo en el primer sentido de ‘información física’: como he dicho, el fisicalismo es una tesis ontológica, no una tesis epistemológica. Más aún, decir que toda la información (correcta) es información física – en el segundo sentido de la noción– es una manera particular de señalar que todo lo que existe es físico.¹⁴

Una vez hecha esta distinción de sentidos, examinemos afirmaciones hechas por Jackson: 1) no existe una cantidad suficiente de información física que incluya ciertos rasgos de las sensaciones corporales; 2) nada que se pueda estructurar en términos de información física captura una experiencia. Si interpretamos estas afirmaciones en el primer sentido de información física, Jackson parecería estar señalando el mismo punto de Nagel; pero como el mismo Nagel reconoce, esto no constituye una prueba de que el fisicalismo sea falso, i.e., que se pueda hacer una interpretación de las mismas afirmaciones en el segundo sentido de información física. Interpretadas estas afirmaciones en el segundo sentido de información física, no permitirían concluir que el fisicalismo es falso sino que estarían presuponiéndolo.

¹⁴ Jackson señala que tenemos todo lo que figuraría en una explicación fisicalista. Pero creo que esto, estrictamente hablando, no es cierto. El esquema fisicalista no es una explicación objetivista de todo en términos de la física, lo cual –pienso– no se puede obtener; el esquema fisicalista es un planteamiento ontológico.

Cuando se introduce el caso de Fred, y se observa que no existe cantidad suficiente de información física acerca del cerebro y del sistema óptico de Fred que nos permita saber qué clase de experiencia tiene cuando ve los dos colores diferentes, no hay duda de que la información física a la que se está haciendo referencia es la que proporciona la neurociencia (es decir, información física en el primer sentido¹⁵). De nuevo, se trataría de una observación de espíritu nageliano.

Si el argumento del conocimiento contiene esta ambigüedad entre dos sentidos de información física, no podemos estar seguros de que la tesis ontológica del fisicalismo haya sido refutada. Reconstruyamos esquemáticamente el argumento para verlo con más claridad:

- *Primera Premisa:* Antes de ser liberada, Mary tiene toda la información física acerca de la visión de color.
- *Segunda Premisa:* Antes de ser liberada, Mary no tiene toda la información acerca de la visión de color (porque aprende algo acerca de ellas al ser liberada).
- *Conclusión:* Mary ha adquirido información que no es física; por tanto, el fisicalismo es falso (pues hay información no-física).

La primera premisa también sería intuitivamente obvia para el fisicalista, pero sólo en el primer sentido de información física: la información que Mary puede adquirir dentro de su habitación, sin duda, es información dada en términos de un lenguaje extensional, objetivista, que presenta sus referentes desde el punto de vista de la tercera persona. Como veremos, el fisicalista intentará problematizar la lectura de la primera premisa bajo el segundo sentido de información física, buscando mantener que, dentro de la habitación, Mary podría tener todo el conocimiento físico en el primer sentido, no en el segundo.

Recordemos que muchos autores han problematizado la segunda premisa (que Mary aprende algo nuevo al salir de la habitación); pero incluso si el fisicalista la aceptara, el argumento del conocimiento refutaría el fisicalismo sólo si este afirmara que toda la

¹⁵ A decir verdad, se trataría de información física en los dos sentidos. El punto es que se trata de información física en el segundo sentido y en el primer sentido.

información correcta acerca del mundo debe estar dada en términos de un lenguaje extensional, objetivista, etc. Pero el fisicalismo no es una tesis acerca de los términos en los que debe formularse nuestro conocimiento del mundo, sino una tesis acerca de la naturaleza de lo que hay en el mundo. De modo que la posición refutada por el Argumento del Conocimiento no es el fisicalismo que se ha estado presentando.

En suma, si interpretamos el Argumento del Conocimiento en el primer sentido de información física, el fisicalista no tendría mayores reparos para aceptarlo, pues no constituiría una amenaza insalvable contra su postura en la medida que no se seguiría la falsedad del fisicalismo (sería más bien un punto de tipo nageliano). De ser así, tanto el caso de Fred como el de Mary permitirían obtener conclusiones de carácter epistemológico y no de carácter ontológico. Pero ¿qué sucede si interpretamos el argumento del conocimiento en el segundo sentido de ‘información física’?

Antes de ocuparme de esta pregunta, quiero hacer algunas observaciones para fortalecer la idea de que el argumento del conocimiento debería leerse bajo el primer sentido de información física. En este orden de ideas, quiero llamar la atención sobre el siguiente pasaje en el que Jackson habla sobre el tipo de conexión que existe entre lo físico y lo fenoménico:

(...) el asunto se complica por la posibilidad de que la conexión entre las cuestiones físicas y los qualia sea como la que algunas veces se dice que se obtiene entre las cualidades estéticas y las naturales. Dos mundos posibles que concuerdan en todos los aspectos “naturales” (incluyendo las experiencias de criaturas sintientes) deben concordar también en todas las cualidades estéticas, pero se dice plausiblemente que las cualidades estéticas no pueden reducirse a lo natural. (Jackson 1982, p. 131)

Si la conexión entre lo fenoménico y lo físico es como la que señala este pasaje, tendríamos una razón para pensar que el enfoque correcto de las premisas que utiliza Jackson no es ontológico sino epistémico (como la que se requeriría para refutar el fisicalismo). La diferencia entre las propiedades estéticas y las naturales no es una diferencia metafísica, y, de manera correspondiente, la diferencia entre propiedades físicas y propiedades fenoménicas tampoco lo sería. Si pensamos en el caso de Mary como un caso acerca de propiedades estéticas y no acerca de propiedades de experiencias, queda claro que no se obtiene una conclusión ontológica: la diferencia

entre las propiedades estéticas y las propiedades físicas no es una diferencia metafísica; de hecho, la existencia de cualquier propiedad estética supone la existencia de propiedades físicas (al punto que podríamos decir que los términos estéticos pueden designar lo mismo que los términos físicos, sólo que lo hacen desde otro modo de concebir, y aunque no podamos reducir explicativamente las propiedades estéticas a propiedades físicas).

Con todo, si el punto de Jackson resulta ser tan cercano al de Nagel –como resultaría de interpretar epistemológicamente las premisas del Argumento del Conocimiento– ¿por qué cree Jackson que su argumento es tan diferente del de Nagel? Aquí vale la pena subrayar que la lectura hecha por Jackson de la posición de Nagel hace parte de las interpretaciones que discutíamos antes en las que se cree que el argumento de Nagel es sobre puntos de vista como instancias y no como tipos.

Jackson piensa que su argumento se diferencia del de Nagel en la medida que su punto no es que “no sabemos cómo es *ser Fred*” sino que “ignoramos una propiedad acerca de su experiencia” (p.132). Pero si interpretamos la pregunta ¿cómo es *ser Fred*? como la pregunta por un punto-de-vista-tipo, vemos que bastaría con añadir a *nuestro* punto-de-vista-tipo el conocimiento de la propiedad experiencial en cuestión para que supiéramos *cómo es ser Fred*.

Si esto es correcto, decir que “ignoramos una propiedad acerca de las experiencias de Fred” implicaría que “no sabemos cómo es *ser Fred*”: no podemos formarnos una idea de su punto-de-vista-tipo, *porque* en nuestra concepción de dicho punto de vista no podemos incluir una propiedad experiencial clave del mismo. Conocer *esa propiedad que ignoramos* nos permitiría *saber cómo es ser Fred*. Esto parece ser exactamente lo que sucede con el murciélago de Nagel: ignoramos un grupo de propiedades acerca de sus experiencias, las propiedades experienciales vinculadas a la percepción del sistema de sonar, por eso no podemos saber cómo es ser un murciélago.

La diferencia que Jackson cree que existe entre su argumento y el de Nagel sólo tiene sentido si pensamos que aquello que es *ser Fred* es un punto-de-vista-instancia y no un punto-de-vista-tipo: incluso si no ignoráramos la mentada propiedad experiencial, no podríamos conocer el punto-de-vista-instancia de Fred. Es verdad que no existe cantidad suficiente de información de ninguna clase que equivalga a “un conocimiento de Fred desde adentro”, es decir, a un conocimiento de su punto-de-vista-instancia. Nagel, por

supuesto, rechazaría una lectura de su argumento en términos de punto-de-vista-
instancia y no punto-de-vista-tipo; lo cual desvanece la supuesta diferencia entre el
argumento de Nagel y el de Jackson.^{16 17}

Jackson subraya que la cualidad especial de la experiencia de Fred que ignoramos y que
es un hecho acerca de ella, queda por fuera de la explicación de la ciencia física porque
no existe suficiente información física –en el primer sentido de la noción– que nos la
explique. De acuerdo a lo que he argumentado, está apuntando a lo mismo que apunta
Nagel con su argumento del murciélago, con la única diferencia de que Nagel pone
explícitamente sus afirmaciones en términos de la distinción objetivo/subjetivo.

Por último, recordemos que Jackson empieza su discusión sobre el epifenomenalismo
con la pregunta de si realmente hay una razón para rechazarlo. Pero creo que la
pregunta es más bien si realmente tenemos una razón para aceptar una postura tan
altamente contraintuitiva. Para ver esto con más claridad, examinemos la defensa que
Jackson hace del epifenomenalismo al final del artículo.

Las razones 1 y 3 parecen ir en contravía de consideraciones metodológicas como el
principio de simplicidad o Navaja de Ockham: proponer una causa con dos efectos es
mucho menos parsimonioso que aceptar una sola causa con un solo efecto. Además,
este cambio en la noción de causalidad que propone Jackson es profundamente
incompatible con nuestra experiencia cotidiana, ¿qué podría resultarnos más obvio que el
hecho de que, por ejemplo, experimentar dolor es lo que me hace llorar?

Con respecto a la segunda razón (el ejemplo del oso polar), podemos ver que todos los
subproductos de los rasgos evolutivos pueden predecirse a partir de una consideración
de sus rasgos físicos, y que todos interactúan causalmente. Así, mientras que la pesadez
de la espesa capa de piel y pelo del oso polar se deriva directamente de las propiedades

¹⁶ Si pensamos no en ‘cómo es ser un organismo completo’ sino en ‘cómo es estar en uno de sus
estados’, se ve claramente que el punto de vista de Nagel es acerca de tipos y no de instancias.
Cuando decimos que alguien que no puede sentir dolor no sabe cómo es ser uno de nosotros, no
decimos que no puede sentir *este* dolor que tengo en este momento, sino que no puede tener un
tipo de estado que nosotros sí podemos tener. Para Nagel, yo podría tener el mismo punto-de-
vista-tipo de un organismo sin ser *ese* organismo, sino siendo el mismo *tipo* de organismo.

¹⁷ Desde luego, tanto el argumento de Nagel como el de Jackson, establecen contundentemente
la misma cuestión epistemológica: la información proporcionada por la ciencia física no nos
permite conocer ninguna propiedad experiencial (como las que aparecen en el punto de vista del
murciélago).

físicas que le permiten dar calor, con la experiencia consciente la situación es completamente diferente, pues las leyes físicas que rigen los estados cerebrales no explican en ninguna forma por qué dan lugar a experiencias conscientes. De esta manera, el argumento de los subproductos epifenoménicos de rasgos evolutivos pareciera colapsar y perder aplicabilidad al caso de la experiencia consciente. Teniendo esto en cuenta reparos, no pareciera haber razones suficientes para abrazar una concepción como la epifenoménica.

Las afirmaciones finales que Jackson hace en el artículo se relacionan con el excesivo optimismo epistemológico que debe tener el fisicalista si cree que rudimentos conceptuales como aquellos de los que se vale la ciencia física son suficientes para entenderlo todo. Presenta una analogía con las babosas marinas para argumentar que sólo entendemos lo que necesitamos entender para sobrevivir y que es muy probable que haya una parte del esquema completo de las cosas que nunca lleguemos a conocer o comprender por más que evolucionemos. Como vemos, este punto del final del artículo tiene un ánimo muy nageliano, pero que Nagel no concibe como un argumento contra el fisicalismo.

1.3.2 La Versión Fuerte

Hace un momento quedó abierta la pregunta ¿qué sucede si interpretamos el argumento del conocimiento en el segundo sentido de ‘información física’? Como señalé hace un momento, la ambigüedad entre los dos sentidos de información física, fue señalada por primera vez por Terence Horgan (1984). En palabras de Horgan, “información física” puede interpretarse (a) en el sentido de ‘información física explícita’ (un enunciado expresa información física explícita sólo si pertenece o se sigue de la teoría física) o (b) en el sentido de ‘información ontológicamente física’ (un enunciado expresa información ontológicamente física sólo si (i) todas las entidades referidas en él son físicas, y (ii) todas las propiedades y relaciones expresadas por los predicados del enunciado son propiedades y relaciones físicas). Sin embargo, Horgan va más allá de simplemente denunciar la ambigüedad y subraya que, en consecuencia, el argumento del conocimiento puede reformularse de dos maneras diferentes:

(1a) Mary tiene *todo el conocimiento físico explícito* de los hechos de la visión de color antes de ser liberada.

(2a) Pero hay cierta *clase de conocimiento* de los hechos acerca de la visión de color que ella no tiene antes de ser liberada.

Por tanto,

(3a) Hay algunas *clase de conocimiento* de los hechos acerca de la visión de color que no es *conocimiento físico explícito*.

(1b) Mary tiene toda la *información ontológicamente física* (es decir, conoce todos los *hechos físicos*) acerca de la visión de color antes de ser liberada.

(2b) Pero hay *ciertos hechos* acerca de la visión de color que ella no conoce antes de ser liberada (porque aprende algo sobre esto al salir de la habitación).

Por tanto,

(3b) Hay *hechos no-físicos* acerca de la visión de color.

Hasta ahora hemos considerado una lectura tipo *a* del argumento de Jackson, que resulta mucho más débil que una lectura de tipo *b*, pues permite una réplica como la que he hecho en el apartado anterior. El asunto es bien distinto si consideramos la versión fuerte de argumento. Si el fisicalismo es verdad, los hechos físicos son los únicos hechos; no obstante, si Mary aprende todos los hechos físicos acerca de la visión de color en su habitación en blanco y negro y, con todo, aprende algo nuevo al abandonarla, ha adquirido un hecho que no es un hecho físico y el fisicalismo no puede ser verdadero.

¿Puede socavarse la formulación fuerte del argumento de Jackson? Varios teóricos han concedido la primera premisa, por lo que conjeturan que la única forma de minar el argumento es atacando la premisa (2b). Desde luego, esto puede hacerse de varias maneras. Por ejemplo, aseverando que Mary no aprende nada nuevo al dejar la habitación. Esta es una posición muy a menudo asociada con Dennett (1991). De acuerdo con Dennett, Mary no aprende nada nuevo al dejar la habitación porque, antes de dejarla, ella ya es capaz de identificar los colores que está a punto de experimentar al hacer complejas inferencias a partir de observaciones de los estados físicos que inducen. La pregunta que ensombrece esta propuesta es ¿el resultado de esta serie de complejas inferencias es el mismo que el experimentar el color? ¿Se trata del mismo carácter cualitativo?

Otra manera de negar la premisa (2b) sería alegar que lo que Mary aprende no es un hecho, sino una habilidad. Según Nemirov (1980, 1990) y Lewis (1988) lo que Mary adquiere es sólo una habilidad para imaginar, reconocer e identificar colores como verde o rojo. Su situación es como la de la persona que conoce todos los hechos sobre bicicletas y sobre cómo funcionan pero nunca ha montado una y aprende algo nuevo (una nueva habilidad o 'saber-cómo') cuando aprende a montar bicicleta. La principal dificultad que se le ha señalado a de esta postura es que parece probable que Mary adquiriera más que una simple habilidad cuando experimenta el rojo por primera vez, parece adquirir también un conocimiento: el de cómo es el rojo en tanto experimentado (cf. Chalmers 1996, Lycan 1995).

Una tercera manera de negar la premisa (2b) es señalar que lo que Mary aprende no es un nuevo hecho sino un viejo hecho de una nueva manera. Según este enfoque, ella es como la persona que sabe que Mark Twain escribió Tom Sawyer y luego aprende que Samuel Clemens escribió Tom Sawyer.

Sin embargo, alguien en esta posición aún aprende algo nuevo, a saber, que Mark Twain y Samuel Clemens son la misma persona. Incluso si el rojo experimentado sólo es el evento cerebral X (ya conocido por Mary), ella aún aprende cosas nuevas cuando experimenta el rojo: por ejemplo, que se puede saber más sobre ver rojo de lo que se conoce cuando está en términos del evento cerebral X.

Una línea de argumentación semejante a esta última es la que indica que el nuevo conocimiento de Mary es conocimiento por aprehensión (by acquaintance), en oposición a conocimiento proposicional que obtuvo en la habitación. Así, no hay razón para creer que adquirir tal conocimiento representa un desafío para el fisicalismo, pues el conocimiento por aprehensión puede ser conocimiento de algo físico, de lo que Mary ya tiene conocimiento proposicional y que no abarca ningún nuevo hecho. Sin embargo, esta no es una forma de negar la premisa B, pues parte de la aceptación de que *hay* un nuevo conocimiento (por aprehensión).

La sensación que nos queda de que la premisa (2b) no ha sido rechazada satisfactoriamente tiene que ver con el hecho de que una de las intuiciones más fuertes

de las que se vale el argumento de Jackson es la que nos convence de que Mary *aprende o adquiere algo nuevo* luego de ser liberada.¹⁸

Desistamos por ahora en nuestro intento de minar la premisa (2b) y examinemos entonces la solidez de la premisa (1b).

Si el conocimiento ontológicamente físico pudiera dividirse en (α) proposicional y (β) por aprehensión (by acquaintance) ¿no podríamos proponer más bien que dentro de la habitación en blanco y negro se puede obtener todo el conocimiento ontológicamente físico proposicional pero no todo el conocimiento ontológicamente físico por aprehensión? Sin argumentos para descartar esta posibilidad, sin tenerla en cuenta, no parece que podamos dar el paso a afirmar que *todo* el conocimiento ontológicamente físico puede ser adquirido dentro de la habitación de Mary; si lo hacemos, presuponemos que cualquier conocimiento que no pueda ser adquirido dentro de ella es no-físico (i.e., que el fisicalismo es falso).

Dicho de otro modo, al formular de manera fuerte el argumento del conocimiento y afirmar que *todo el conocimiento ontológicamente físico* puede adquirirse dentro de la habitación –ignorando la mencionada posibilidad–, efectivamente, por implicación, afirmamos que cualquier conocimiento que no puede ser adquirido dentro de ella es ontológicamente no-físico, y *presuponemos* que el fisicalismo es falso (predeterminando por ende el resultado del experimento mental en favor de una refutación del fisicalismo). Luego de este paso, el argumento sería reforzado por el poder de la intuición de la segunda premisa, que nos indica que Mary aprende o adquiere algo nuevo al salir de la habitación. Pero la apariencia de inferencia sería ficticia, habría una especie de petición de principio.

¹⁸ Además de la intuición vinculada al argumento del conocimiento, Jackson alude a otra intuición que subyace a varios argumentos anti-fisicalistas, a saber, la intuición según la cual hay un mundo posible con organismos físicamente como nosotros pero que no tienen conciencia. Creo que los problemas del argumento modal y del argumento del conocimiento empiezan por el hecho de que algunos autores discuten sinceramente las intuiciones que subyacen a ellos; por ejemplo, en el caso de la intuición modal, se ha propuesto sinceramente que es contraintuitivo que pueda haber zombis. Esto evidentemente problematiza y relativiza lo que significa decir que algo es “intuitivamente una verdad obvia”: el poder de una intuición dependería de si comparto o no comparto la intuición.

En la medida en que estamos dispuestos a admitir que Mary aprenderá algo nuevo al dejar su habitación, no podemos simplemente admitir que ella ha adquirido todo el conocimiento ontológicamente físico cuando estaba dentro de la habitación. Una ilustración puede ayudar a mostrar más este punto. Imaginemos a Mary, con visión normal, encerrada esta vez en una habitación llena de cubos de colores. Recolecta los 99 cubos rojos que tiene a la vista y concluye que la afirmación “todos los cubos en la habitación son rojos” es verdadera, aunque sabe que podría haber otro cubo en un armario de la habitación, el armario está con llave y no sabe el color del cubo. Mary quiere saber si la conclusión que formuló todavía se mantiene. Ella podría pensar que ya tiene todos los cubos rojos, por lo cual el cubo que queda debe ser un cubo no-rojo y la afirmación “todos los cubos en la habitación son rojos” es falsa. Sin embargo, obviamente estaría sesgando su razonamiento en favor de una refutación de la mencionada conclusión. Afirmer que los cubos que ella tiene son todos los cubos rojos sabiendo que hay otro cubo en el armario que ella no ha visto, es definir el cubo que queda como no-rojo (asegurando que la conclusión previa refutada).

En suma, podría haber un caso en el que la primera premisa del argumento en la versión *b* sería falsa, caso en el cual el argumento no se seguiría (sea o no “*intuitivamente una verdad obvia*”): Mary podría no tener todo el conocimiento ontológicamente físico, podría no conocer todos los hechos físicos. Ya he mencionado que el fisicalista no tiene que comprometerse con que todos los hechos físicos se conocen mediante el vocabulario de la ciencia física (i.e., puede admitir que hay hechos físicos que no es posible conocer únicamente con el vocabulario de la ciencia física).

2. EL DEBATE ALREDEDOR DE LOS ARGUMENTOS CONTRA EL FISICALISMO

A continuación se examinan las consideraciones que, sobre el problema que se viene rastreando en la tesis, han hecho algunos teóricos como Levine (1983, 1993), Chalmers (1995), Loar (1998), Papineau (1998). En este sentido, la discusión contemporánea sobre el fisicalismo acerca de la conciencia fenoménica ha incluido algunos elementos aportados por estos filósofos, como la noción de 'brecha explicativa' de Levine (junto con el preludio a la teoría de los conceptos fenoménicos que hace), así como el *problema difícil* de Chalmers.

2.1 LEVINE: La Brecha Explicativa

Como señalé en las páginas anteriores, la extensa y meticulosa argumentación desarrollada por Joseph Levine (1983) aparentemente es una línea que le podría interesar al fisicalista si desea preservar la integridad de su tesis ante los argumentos modales de Kripke. La argumentación de Levine está concebida básicamente como un intento de transformar el argumento ontológico de Kripke en un argumento epistemológico. De acuerdo con Levine, el argumento de Kripke está construido sobre dos afirmaciones clave: (1) que los enunciados de identidad que usan designadores rígidos en ambos lados son necesarios (i.e., verdaderos en todos los mundos posibles) si son verdaderos; (2) es concebible que los enunciados de identidad psicofísicos sean falsos en algún mundo posible (por lo tanto no son necesarios y tampoco verdaderos). La idea de Levine es que (2) está basada en una intuición insuficiente para dar sustento a una tesis metafísica y que más bien sustenta una tesis epistemológica. Con todo, Levine enfatiza que si bien de su reconstrucción del argumento no se sigue que el fisicalismo sea falso, sí se sigue un problema diferente para el enfoque fisicalista.

Recordemos que, de acuerdo con Kripke, la aparente contingencia (la ‘ilusión de contingencia’) de ciertos enunciados de identidad desaparece en casos como “*calor=movimiento molecular*”, pero no en el caso de la identidad “*dolor=activación de las fibras C*”; en consecuencia, la verdad de este segundo enunciado quedaba en duda: bien podría ser que la contingencia no fuera ilusoria. Causa de lo anterior era que en el caso del segundo enunciado de identidad no era posible aplicar el modelo que aplicamos en el primer caso para disipar dicha ilusión de contingencia (el modelo que he llamado ‘de identificación por medio de una propiedad contingente asociada’).

En este sentido, Levine analiza algunas de las respuestas que el fisicalista tiene a mano. Tenemos, para empezar, la respuesta funcionalista, según la cual la intuición de que el dolor podría existir sin la activación de las fibras C se explica en términos de la múltiple realizabilidad de los estados mentales. Parece imaginable, sin embargo, que en algún mundo posible el enunciado de identidad psicofuncional sea falso (como en el caso de Block (1980) –en el mundo actual– de la nación china que parece posible que realice el estado funcional F, aunque no sienta dolor).

De acuerdo con Levine, algunos funcionalistas han tratado de replicar que lo que parece indicar el ejemplo de Block es que las descripciones funcionales son simplemente demasiado abstractas como para capturar los rasgos esenciales de las experiencias sensoriales cualitativas. Estos funcionalistas argumentan que aunque es suficiente para ser consciente que una entidad realice la descripción funcional apropiada, la forma particular en la que es experimentado un estado cualitativo estaría determinada por la naturaleza de la realización física. Sin embargo, Levine señala que si el funcionalista adopta esta posición (de replegarse ligeramente hacia el fisicalismo), podría aplicarse de nuevo argumento original de Kripke: podría imaginarme esa forma particular de experimentar el estado cualitativo sin que se realice el estado funcional (o sin determinada realización física).

Así pues, de acuerdo con Levine, el dilema que en el que Kripke deja al funcionalismo es el siguiente: o bien se mantiene en que –a pesar de la intuición– los enunciados de identidad psicofuncionales no son contingentes, o bien debe abandonar sus pretensiones teóricas.

Y bien, ¿cuál es el problema con la primera opción? De acuerdo con Levine, el mismo Kripke le da todas las herramientas que requiere el fisicalista o el funcionalista para responder al desafío que él mismo les ha puesto, al enfatizar el golfo entre la posibilidad epistemológica y la posibilidad metafísica: lo que intuitivamente parece ser el caso es un asunto meramente epistemológico, no ontológico; y dado que la posibilidad epistemológica (i.e., concebibilidad) no sería suficiente para la posibilidad metafísica, no deberíamos darle mucha importancia a lo que sea intuitivamente contingente, pues muchas veces puede resultar siendo metafísicamente necesario.

2.1.1 ¿Una *mera* cuestión epistemológica?

Con todo, Levine subraya que, si bien se trata de un asunto claramente epistemológico, no se puede simplemente descartar el problema diciendo que se trata de una *mera* cuestión epistemológica.

Hemos visto que enunciados como “*dolor=activación de las fibras C*” (o “*dolor=estado funcional F*”) comparten una diferencia importante con el enunciado “*calor=movimiento de las partículas*”, a saber, que en ellos la ilusión de contingencia no se puede disipar (al menos de la manera *estándar*) mientras que en aquél sí. Para Levine esta diferencia se debe a otra diferencia subyacente a estos enunciados: mientras que el tercero es un ejemplo de una identidad ‘completamente explicativa’ (en la cual podemos ver que nada importante ha quedado por fuera –razón por la cual esperaríamos una explicación que disipara cualquier aparente contingencia), los otros dos parecieran dejar algo crucial por fuera (hay un *vacío* en su pretensión explicativa; por lo cual no vemos cómo descartar su aparente contingencia). Para Levine es este *vacío* el que da lugar a las objeciones tipo Kripke.

¿Qué significa que el tercer ejemplo sea ‘completamente explicativo’? Quiere decir que nuestro conocimiento de química y física hace inteligible cómo es que algo como el movimiento de las moléculas puede jugar el rol causal que asociamos con el calor: Levine indica que una vez entendemos cómo es efectuado este rol causal por el movimiento de las moléculas, no hay nada más que necesitemos entender.

Levine reconoce que en los otros dos casos el rol causal del dolor también ha sido explicado (se ha explicado un aspecto importante de lo que hay que explicar acerca del dolor); pero nuestro concepto de dolor incluye algo más que su rol causal: involucra su

carácter cualitativo, el 'cómo se siente'. Este carácter queda sin explicar al descubrir que la activación de las fibras C cumple este rol causal, *¿por qué el dolor debería sentirse como se siente?* O, en otras palabras *¿Cómo se sigue del hecho de que el dolor sea activación de las fibras C el hecho de que tenga el carácter cualitativo que tiene?*

Levine afirma que en este sentido no pareciera haber nada en la activación de las fibras C que la ajuste *naturalmente* (*¿evidentemente? ¿Obviamente?*) al carácter cualitativo del dolor, más de lo que se ajustaría a cualquier otro conjunto de propiedades fenoménicas; en consecuencia, la identificación del lado cualitativo del dolor con la activación de las fibras C deja en el completo misterio la conexión entre ambos.

El punto que está señalando Levine puede ilustrarse más aún de la siguiente manera: supongamos que R y V son las descripciones en términos de la física de ver rojo y verde, respectivamente. Cuando consideramos el carácter cualitativo de nuestras experiencias visuales, la diferencia no se explica al apelar a R y a V, pues no explican realmente por qué tengo cierto tipo de experiencia cualitativa y no otra (por eso podemos imaginar V o R subyaciendo a la experiencia asociada de hecho con R y viceversa).¹⁹

2.1.2 La Brecha Explicativa

Al comienzo de este apartado señalé que, de acuerdo con Levine, el argumento modal de Kripke se basa en una intuición insuficiente para sustentar una tesis metafísica. Esta intuición, no obstante, es suficiente para apoyar una tesis epistemológica, a saber, la tesis de que los enunciados de identidad psicofísicos dejan una *brecha explicativa*: tales enunciados tienen un *vacío* en su pretensión explicativa, dejan algo crucial por fuera (la relación entre el lado fenoménico de la identidad y el lado físico de la misma).

Lo primero que anota Levine es que esta noción de *brecha explicativa* será tan clara como puede serlo la noción correspondiente de *explicación*. De acuerdo con el modelo

¹⁹ Levine advierte que recurrir aquí a identificar los qualia con sus roles funcionales es de ninguna ayuda, pues: 1) recordemos que para algunos autores un estado funcional se sentiría de tal o cual manera *en virtud de su realización física*; 2) sigue cabiendo la pregunta de por qué el tipo de estado que realiza la función del dolor debería sentirse de la manera que lo hace. (Nótese que la pregunta análoga con respecto al calor no es tan poderosa – *¿por qué el movimiento de las partículas cumple el rol funcional del calor?*–, una comprensión de química y física es todo lo que se necesita para responder esta pregunta).

de explicación deductivo-nomológico, un evento particular *e* es explicado cuando se muestra que es deducible de leyes generales junto con una descripción relevante de la situación particular. Pero Levine resalta que los enunciados de identidad fisicalista y funcionalista obviamente podrían ser empleados como premisas en una deducción concerniente a los estados psicológicos de alguien (además, soportan contrafácticos y podrían usarse –junto con otras premisas– para deducir toda clase de particulares). Así, pues, el modelo deductivo-nomológico de explicación sería insuficiente para aclarar la noción *brecha explicativa*, en la medida que no le resulta problemático que la relación entre el lado fenoménico y el lado físico de la identidad estén en la oscuridad.

Levine puntualiza entonces que la noción de explicación que se requiere para aclarar la noción de *brecha explicativa* debe dar cuenta de lo que significa hacer que un fenómeno se vuelva *inteligible*, y debe incluir reglas que determinen cuándo es inapropiada la demanda de mayor *inteligibilidad*. De este modo, en una explicación satisfactoria, es decir, en una que cumpla el criterio de inteligibilidad, nada parece ser “dejado por fuera”, no se nos deja con la sensación de que algo que debería caer bajo la explicación no ha sido incluido.

Si bien no da un tratamiento teórico más extenso a la noción de inteligibilidad, Levine piensa que hay razones para buscar una explicación de ese tipo: por un lado, la conciencia sólo parece surgir en el nivel de organización macroscópico (lo cual debería esperarse si fuera un asunto de organización funcional compleja); por otro lado, el fisicalismo –según Levine– implica que para todo fenómeno que no sea descriptible en términos de las magnitudes físicas fundamentales, existe un mecanismo que es descriptible en términos de las magnitudes físicas fundamentales (de manera tal que las ocurrencias de este son inteligibles en términos de ocurrencias de aquellas). Esta idea implica que hechos primitivos no deberían surgir en el dominio de lo macroscópico, por lo cual parecería extraño que hechos primitivos como el dolor surjan en este nivel de organización.

El asunto de la brecha, señala Levine, puede formularse entonces de la siguiente manera: *si* no hay nada en la activación de las fibras C que explique por qué tenerla tiene ese carácter cualitativo, *entonces* inmediatamente se vuelve imaginable que haya activaciones de las fibras C sin sentir dolor y *viceversa*. No puede decirse lo mismo en el

caso de “*calor=mp*”: cualquier rasgo que haya que explicar del calor es explicado por el movimiento molecular. Nada se queda por fuera.

La existencia de una brecha explicativa como la que se ha descrito, siguiendo a Levine, explica la aparente contingencia de los enunciados de identidad fisicalista y funcionalista, y el fracaso al intentar desvanecer su aparente contingencia de la manera estándar (es decir, mediante el modelo propuesto por Kripke). En palabras de Levine: “*es la no-inteligibilidad de la conexión entre el sentimiento de dolor y su correlato físico, la que subyace a la contingencia aparente de esta conexión*” (Levine 1983, p. 359).

Pero para Levine el problema podría ir más allá: de la existencia de la brecha se seguiría que incluso si algunos enunciados de identidad psicofísicos son verdaderos, no podemos determinar exactamente cuáles son verdaderos. El asunto es el siguiente: decir que el dolor es idéntico a un estado físico no nos permite determinar qué medida de semejanza/desemejanza física usar. Si un organismo tiene estados físicos ligeramente diferentes a los nuestros ¿tiene o no nuestros mismos estados cualitativos? ¿Tiene estados cualitativos en absoluto? El hecho de que el dolor sea una clase de estado físico, no nos dice qué tan gruesa o delgadamente cortar nuestras clases físicas cuando se determina a qué estado físico es idéntico. El problema persiste incluso si nos trasladamos a un marco funcionalista: si creemos que el dolor es idéntico a un estado funcional, ¿qué medida de semejanza/desemejanza funcional usamos al juzgar si una criatura tiene o no nuestros estados cualitativos?

Si fuera cierto que tener una constitución física/funcional diferente es suficiente para tener diferentes estados cualitativos (o incluso para carecer de ellos), entonces debe haber algo que nos indique *qué tan diferente* debe ser esta constitución. Si hubiera alguna conexión intrínseca inteligible entre la activación de las fibras C (o el estado funcional F) y el estado cualitativo en cuestión –tal que experimentar el segundo fuera inteligible en términos de las propiedades del primero–, podríamos derivar fácilmente la medida de semejanza. Si contáramos con esta conexión, podríamos hacer un razonamiento como el siguiente: la activación de las fibras C tiene las propiedades α , β , γ , etc., por lo tanto debe sentirse de tal o cual manera. Cualesquiera propiedades de la activación de las fibras C (o del estado funcional F) que explicaran la sensación de dolor

determinarían las propiedades que un estado físico/funcional debe tener para contar como sensación de dolor. Pero esa es precisamente la brecha que no se ha llenado.

2.1.3 El Alcance de la Brecha y el asunto de la Inteligibilidad

Hace un momento veíamos que, de acuerdo con Levine, nuestros conocimientos de química y de física nos permiten ver cómo el movimiento de las moléculas puede jugar el rol causal del calor: una vez entendemos cómo este rol causal es efectuado por el movimiento de las moléculas no sentimos que haya algo más que necesitemos entender acerca de esta identidad (vemos claramente la conexión entre las dos cosas). Sin embargo, en el caso de los enunciados de identidad psicofísicos no podemos ver cómo las características del estado físico/funcional satisfacen (todas) las características del estado cualitativo; es más, no hay *nada* en la activación de las fibras C que haga que se ajusten a las propiedades fenoménicas del dolor (más que a cualesquiera otras propiedades fenoménicas). De lo anterior se sigue, para Levine, que identificar el lado cualitativo del dolor con la activación de las fibras C deja *en el completo misterio* la conexión entre esta y eso con lo que lo identificamos; no puede verse *ninguna* relación entre ambas.

Creo que la mayor parte de este argumento es incontestable. En el caso de los enunciados de identidad psicofísicos o psicofuncionales, no se ve la relación explicativa entre el lado físico/funcional de la identidad y el lado fenoménico. *Prima facie*, hablar de un estado cerebral no es hablar del carácter cualitativo de una experiencia, y tampoco lo implica. Y si esta fuera la única posible relación entre los dos ítems, estarían completamente disociados: no se podría ver cuál es su vínculo.

Pero es aquí donde empieza mi disenso con Levine, relacionado con el alcance que él piensa que tiene su argumento. En particular, cuando dice que “*la identificación del lado cualitativo del dolor con la activación de las fibras C (...) deja la conexión entre esta y aquello con lo que lo identificamos en el completo misterio*” (Levine 1983, p. 357), creo que va más allá de lo que el argumento permite afirmar. Una cosa es que (*prima facie*) hablar de un estado cerebral no sea lo mismo que hablar del carácter cualitativo de una experiencia, o que el primero no implique el segundo, otra cosa es que no se vea *ninguna* relación entre los dos ítems.

Sin embargo, si se admite –como he admitido– que no se ve la relación explicativa entre el *explanandum* y el *explanans* propuesto (es más, si estamos de acuerdo es que posiblemente no vamos a llegar a ver la relación explicativa entre ambos, pues hacen parte de dos marcos epistémicos o modos de concebir diferentes –uno objetivo y otro subjetivo–) ¿no queda la conexión entre el *explanandum* y el *explanans* propuesto en el completo misterio? ¿No quedamos impedidos para establecer una conexión entre ambos por la falta de este vínculo explicativo “natural” del que habla Levine? Creo que no. Creo que Levine se equivoca al sugerir que la *relación de explicación* entre los dos ítems es la única relación posible entre ambos. De la aceptación de que existe una brecha *explicativa* (epistemológica, insalvable, entre el ‘modo de concebir objetivo’ y el ‘modo de concebir subjetivo’) *no se sigue* la afirmación de que no podemos ver conexión alguna posible entre los dos ítems de la identidad psicofísica. Como he señalado antes, somos llevados a la conclusión de que el enunciado de identidad psicofísico es un hecho por razones como el argumento causal, o lo que Churchland llama el ‘argumento de la *dependencia neuronal* de todos los fenómenos mentales conocidos’, etc. En este sentido, si bien hay algo que no entendemos de la mentada conexión, no es verdad que quede “en el completo misterio”.²⁰

Podemos preguntar entonces *¿tiene que darse* este vínculo explicativo “natural” entre los dos ítems para que podamos pensar que tienen una conexión como la señalada en el enunciado de identidad psicofísico? Al parecer no. De ser así, el enunciado de identidad psicofísico no tiene que ser completa y perfectamente explicativo –como de hecho sucede– para que estemos autorizados a pensar que es verdadero.

No perdamos de vista que la cuestión de la identidad es una cuestión ontológica (en las que, recordando a Frege, se señala que dos modos de designación apuntan al mismo objeto), por lo cual la identidad no tiene que ser completa y perfectamente explicativa, lo

²⁰ Como se puede ver en enfoques como el de Nagel o el de McGinn, es posible que tampoco estemos autorizados a decir que “no hay *nada* en la activación de las fibras C que haga que se ajusten a las propiedades fenoménicas del dolor (más que a cualesquiera otras propiedades fenoménicas).” Bien podría resultar que seamos nosotros, los seres humanos –con todas las limitaciones propias de nuestro aparato cognitivo-, quienes no podamos ver semejante cosa (quizás otros seres inteligentes, mucho más que nosotros, puedan encontrar la relación explicativa entre los dos ítems de la identidad psicofísica).

cual es un asunto epistemológico. De modo semejante a lo que señalará Papineau, exigir que las identidades sean completa y perfectamente explicativas para que podamos admitirlas estaría fuera de lugar.²¹ Puede que los enunciados de identidad psicofísicos o psicofuncionales no impliquen nada acerca del carácter cualitativo de las experiencias conscientes, pero no es verdad que la conexión entre los estados físicos/funcionales y las experiencias conscientes sea ‘completamente misteriosa’ (si bien no es del todo clara).

Ahora bien, más adelante Levine sugiere que en una explicación satisfactoria (i.e., donde un fenómeno se vuelve *inteligible*) nada parece ser “dejado por fuera” en el sentido de que no quedamos con la sensación de que algo que debería explicarse no es explicado. Debería resultarnos inquietante que se vincule de una manera tan estrecha la noción de inteligibilidad (y, por tanto, la de explicación) con la presencia o ausencia de una “sensación de que algo que debería explicarse no lo es” (Levine 1983, p. 358). Lo que se esperaría es que la inteligibilidad vaya más allá de una mera sensación intuitiva, pues si se reduce simplemente a eso ¿qué nos impediría descartar este palpito como algo originado, por ejemplo, en la ignorancia de parte de quien tiene la sensación, como algo que tiene que ver más con un sujeto que con la teoría misma? El poder de la argumentación de Levine parecía radicar en que apuntaba a un problema conceptual en la explicación misma (que ningún sujeto competente podría resolver). Pero el tratamiento que Levine hace de la brecha explicativa podría debilitarse significativamente al poner a depender la noción de inteligibilidad (y con ella, la de explicación) de algo tan impreciso y vago como una sensación.

En las líneas anteriores Levine había presentado el asunto de la brecha de una manera mucho más precisa y poderosa, esto es, como el problema *de principio* de que no se pueda hacer una derivación de los hechos subjetivos a partir de hechos neurofisiológicos. De este modo, algo era explicado (o ‘hecho inteligible’) cuando ‘no quedaba nada por fuera’ en el sentido de que se podía hacer una derivación completa del explanandum a

²¹ En lugar de identidades completa y perfectamente explicativas, lo que se debería exigir son las razones por las que creemos que se trata de una identidad verdadera.

partir del explanans –más que cuando no se presentaba una sensación de que algo que debería explicarse no es explicado–.

Ahora parece haber un retroceso, pues la relación con la mencionada sensación hace de la inteligibilidad una noción más bien confusa/difusa. Veamos por qué. Supongamos que la noción de inteligibilidad depende de la sensación de que algo que debería explicarse no es explicado. ¿Qué sucedería en el caso de que alguien manifieste tener esta sensación ante la identidad “*calor=mp*”? (¿No parece posible que esto suceda cuando decimos que *todo* lo que hay que explicar acerca del calor es explicado por el movimiento molecular?) Si sucediera masivamente ¿deberíamos decir que hay una brecha entre el calor y el movimiento molecular? ¿Cómo establecemos que este caso es diferente al de aquel que manifiesta tener la sensación ante la identidad “*dolor=activación de las fibras C*”? ¿Cuándo deberíamos tener en cuenta esta sensación y cuándo no? La respuesta que encontramos en Levine a este tipo de interrogantes es muy poco satisfactoria, en especial si tenemos en cuenta que se trata de un punto central para su argumentación:

“Con todo, me han dicho que el valor de G , la constante gravitacional, no se deriva de ninguna de las leyes básicas. Es un hecho dado, primitivo, acerca del universo. ¿Esto nos deja con una sensación de que algo que debería ser explicado no lo es? ¿O esperamos que algunos hechos de la naturaleza debieran aparecer arbitrariamente de esta manera? Me inclino a tomar la segunda actitud con respecto a G . Así, podría preguntarse, ¿por qué la conexión entre como qué es estar en un estado funcional (o físico) particular y el estado mismo requiere una explicación para hacerse inteligible? Sin una explicación teórica de la noción de inteligibilidad que tengo en mente, no puedo proporcionar una respuesta realmente adecuada a esta pregunta.” (Levine 1983, p. 358)

Este pasaje deja “*en el completo misterio*” el papel que juega la aludida sensación: ni siquiera está claro cuándo se presenta, o si debiera tenérsela en cuenta cada vez que se presenta; por ende, no queda claro que la sensación sea relevante para establecer una posible insuficiencia explicativa, una brecha. Sin lo anterior, es insuficiente –acaso

irrelevante— la presencia de dicha sensación para anteponer la exigencia de inteligibilidad a la conexión entre un estado físico (o funcional) particular y un estado cualitativo.²²

Un par de observaciones más alrededor de la argumentación de Levine. Primera. Levine señala que una posible formulación del asunto de la brecha es el siguiente condicional: *si* (1) no hay nada en la activación de las fibras C que explique por qué tenerla tiene el carácter cualitativo que tiene (dicho de otro modo, *si* la manera particular en que se siente tener una experiencia no se explica —o se hace inteligible— al entender las propiedades físicas o funcionales del estado físico propuesto), *entonces* (2) inmediatamente se vuelve imaginable que haya activaciones de las fibras C sin sentir dolor y *viceversa*²³. La duda que surge en este respecto se origina precisamente en el “*viceversa*”, y podría ilustrarse con el siguiente caso: pensemos en Juan, una persona del renacimiento, cuya vida transcurre antes de los trabajos de Dalton sobre la teoría atómica. En lo que a Juan respecta, puede imaginar una piedra que no esté compuesta de átomos, aunque siga teniendo peso, masa, volumen, etc. (puede imaginarlo porque, de hecho, en lo que a él respecta, las piedras no están hechas de átomos). Sin embargo, de lo anterior no se sigue que no haya nada en la teoría atómica que explique las propiedades de la piedra.

Pero no tenemos que ir tan lejos, pues podríamos pensar que Juan es una persona de nuestro tiempo que simplemente ignora que la teoría atómica se aplica a los objetos cotidianos de la realidad: una vez más, Juan podría imaginar una piedra que no está compuesta de átomos. No obstante, pareciera que una vez descubre y se convence de que la teoría atómica se aplica a los objetos cotidianos de la realidad, Juan ya no podría imaginar lo que antes parecía poder imaginar.

Así, otro elemento que evidencia el caso de Juan es que la capacidad de las personas para imaginar situaciones posibles podría variar según los conocimientos que van adquiriendo (véase Churchland 1985). Esto tendría relación con el hecho de que nuestros conceptos puedan cambiar o ser sustituidos con el tiempo, por lo que no sólo que lo que es concebible en un momento (para un sujeto o una época histórica) puede no

²² No me referiré al asunto de la explicación en relación con hechos primitivos del universo, que también hace parte del pasaje (a esto me referiré con más cuidado más adelante, cuando examine los argumentos del Chalmers).

²³ Esto es, si (2) entonces (1): si es imaginable, quiere decir que hay insuficiencia explicativa.

serlo en otro momento subsiguiente, sino que lo concebible no determina lo que puede suceder en el mundo (lo posible): no tiene sentido decir que la identidad “*calor=mp*” se mantiene en t_1 y no en t_2 .

Esta reflexión nos lleva a otra afirmación de Levine, según la cual nada en la activación de las fibras C la ajusta *naturalmente* al carácter cualitativo del dolor (más de lo que se ajustaría a cualquier otra experiencia consciente). Para superar la evidente imprecisión del término “naturalmente”, podríamos interpretarlo de la siguiente manera: α se *ajusta naturalmente* a β cuando se puede notar a primera vista que de las propiedades de β se siguen/derivan las propiedades de α . Esta concepción, que podríamos llamar de ‘ajuste natural *prima facie*’ sucumbiría ante los argumentos recién expuestos: si bien α puede no *ajustarse naturalmente [prima facie]* a β , esto podría deberse a la carencia de algunos conocimientos que, una vez obtenidos lo hacen ajustarse (ya no *prima facie* sino *segunda facie*, si cabe la expresión).

Segunda. Levine anota que la mera identificación de experiencias conscientes con estados físicos no nos da pautas sobre qué tan grueso o qué tan delgado debemos cortar dichas clases físicas (por ejemplo ¿la menor variación en la clase física afecta la identidad o cuál es el rango de variación aceptable?). Ahora, si pudiéramos ver la conexión explicativa entre la activación de las fibras C (o el estado funcional F) y el estado cualitativo en cuestión, tal que experimentar el segundo fuera inteligible en términos de las propiedades del primero, podríamos derivar ese rango de variación. De lo anterior se sigue, prosigue Levine, que la verdad o falsedad de la identidad aunque fuera metafísicamente actual podría ser epistemológicamente inaccesible: no sabríamos exactamente cuáles son los estados (o el rango de estados) que componen la clase física a la que es idéntica la experiencia.

No parece que se pueda decir algo en contra de lo anterior. Sin embargo, este argumento no problematiza la posición del fisicalista –al menos no tal como la estoy concibiendo– más de lo que la problematiza la existencia de la brecha explicativa misma: lo que plantea el fisicalista es que las identidades en cuestión son el caso, independientemente de nuestra capacidad de acceder epistémicamente en mayor o menor medida a ellas. Como señalé en líneas anteriores, en vista de que un vínculo *explicativo* entre ciertos estados cerebrales y las experiencias conscientes no sería la única razón para pensar que hay una conexión entre ambos, el fisicalista podría tener

argumentos independientes para pensar que la identidad es el caso, a pesar de los obstáculos epistemológicos señalados por Levine.

Tercera. Durante su discusión sobre la inteligibilidad de las identidades fisicalistas, Levine señala que la postura fisicalista implica lo siguiente: para todo fenómeno de la realidad que no sea descriptible en términos de las magnitudes físicas fundamentales, existe un mecanismo descriptible en términos de las magnitudes físicas fundamentales, tal que las ocurrencias del primero son inteligibles en términos de ocurrencias del segundo. De acuerdo con Levine, esta idea implica que los hechos brutos (i.e., simples) no deberían surgir en el dominio de lo macroscópico, por lo cual parecería extraño que hechos primitivos brutos/simples como el dolor surjan en este nivel de organización.

Pero el argumento anterior no es forzoso. Los fisicalistas comúnmente sostienen que todos los fenómenos son físicos (o bien 'dependientes ontológicamente de lo físico'), y en esta medida podrían admitir que a todo fenómeno que no sea *descriptible* en términos de las magnitudes físicas fundamentales le corresponde un mecanismo descriptible en términos de las magnitudes físicas fundamentales; lo anterior, no obstante, no compromete al fisicalista con la afirmación de que las ocurrencias del primero tengan que ser inteligibles en términos de ocurrencias del segundo. El fisicalista bien puede admitir la existencia de propiedades emergentes, propiedades que no se pueden reducir a (ni predecir a partir de) su mecanismo subyacente o el comportamiento de sus partes constituyentes (como sucede en los sistemas complejos). Una propiedad emergente es precisamente un caso en el que las ocurrencias de un fenómeno no 'son inteligibles' en términos de las ocurrencias de su mecanismo físico subyacente; más aún, este tipo de propiedades son hechos primitivos brutos/simples que surgen en el nivel de organización macroscópico.

2.1.4 Dos formas de concebir el fisicalismo

Quiero recapitular algunos elementos que la argumentación de Levine parece establecer de manera convincente: en primer lugar, parece una cuestión epistemológica de principio el que no sea posible derivar hechos fenoménicos a partir de hechos físicos. Se sigue (en segundo lugar), que aunque las identidades psicofísicas sean el caso, podríamos no llegar a conocer exactamente cuáles son el caso. Llegamos así (en tercer lugar) a un punto aún más general: que no existe ningún vínculo *explicativo* entre los hechos

fenoménicos y los hechos físicos (esto es lo que significa que sean conceptualmente independientes los unos de los otros); en lugar de este vínculo explicativo, tenemos una *brecha explicativa*.²⁴

Ahora bien, se diría que la reconstrucción que Levine hace del argumento de Kripke le heredaría un problema serio al fisicalismo, a saber, el de la brecha explicativa. No obstante, creo que este resulta un problema serio para una especie particular de fisicalismo, que no es el que he estado discutiendo. Si el fisicalismo incluyera la tesis de que, en la medida en que todo es físico, *debe ser posible* hacer una derivación de cualquier hecho a partir de los hechos descritos en términos de la ciencia física, la tesis del fisicalista estaría en grave peligro. A esta forma de ver el fisicalismo se le podría llamar *fisicalismo epistemológico*.

Llama la atención que Levine no afirme que su argumentación prueba que el fisicalismo es falso. Si el fisicalismo que Levine tiene en mente fuera *fisicalismo epistemológico*, no debería dudar en decir que se ha probado que el fisicalismo es falso. Lo que parece yacer aquí es la idea del fisicalismo, *escindido de cualquier posición epistemológica*, no ha sido puesto en peligro (si bien se probaría la falsedad cualquier especie de fisicalismo que tuviera pretensiones epistemológicas). He insistido en que el fisicalismo planteado aquí es una tesis *ontológica*, tal que no implica ninguna afirmación acerca de nuestras capacidades epistémicas o conceptuales de conocer el mundo. Este es el tipo de fisicalismo que no se ha probado falso de acuerdo con Levine; se le podría llamar (aunque, *stricto sensu*, resulte redundante) *fisicalismo ontológico*.

Por último, resaltemos la diferencia que Levine establece entre su posición y la de Kripke. Mientras que para el segundo el fisicalismo ontológico es falso (basado en la idea de que el fisicalista no puede responder al desafío de explicar la ilusión de contingencia de las identidades psicofísicas), para el primero es el fisicalismo epistemológico el que es falso (pues no es posible derivar hechos fenoménicos a partir de hechos físicos). Ahora, he señalado que en el caso de las identidades psicofísicas no resulta obligatorio usar el modelo que Kripke utiliza para disipar esta ilusión de contingencia de otros casos, pues

²⁴ Esto no significa, desde luego, que la única conexión que puede encontrarse entre ambos ítems sea *explicativa*. He señalado cómo podrían encontrarse buenas razones para pensar que existe un vínculo a partir de consideraciones causales.

se puede formular un modelo alternativo basado en la distinción entre lo que es concebible y lo que es posible (véase mi discusión sobre los argumentos de Kripke). Este modelo alternativo explicaría incluso por qué es predecible la concebibilidad de la negación de la identidad, basado en la idea que se trata de una identidad empírica *a posteriori* y que además involucra dos sistemas conceptuales diferentes.

Una idea sugestiva en este respecto, que será analizada con más cuidado en lo que sigue, es que las experiencias conscientes podrían ser un modo de presentación de ciertos estados y procesos físicos para la primera persona, i.e., para el sujeto que los padece. Esto implicaría que son idénticos, que es una identidad que sólo se puede conocer *a posteriori* y que para cada uno de los idénticos hay un sistema conceptual distinto (lo cual, a su vez, explicaría que pueda imaginarse la falsedad de la conclusión aunque sea metafísicamente un hecho). Más adelante veremos hasta donde es posible llevar esta sugerencia.

2.2 LEVINE: FISCALISMO ONTOLÓGICO Y FISCALISMO EPISTÉMICO

En el apartado anterior señalé que la brecha explicativa constituiría un problema para una especie particular de fiscalismo: aquel que equivale a la tesis que, en la medida en que todo es físico, *se sigue que debe ser posible hacer una derivación de cualquier hecho a partir de los hechos descritos en términos de la ciencia física*. Observé que este *fiscalismo epistémico* no es obligatorio para el fiscalista, ni es el que he estado presentando, y que tampoco tiene una relación necesaria o relevante con éste; he insistido en que el fiscalismo es una tesis ontológica que no implica ninguna afirmación acerca de nuestras capacidades epistémicas o conceptuales de conocer el mundo. También señalé que en la medida que Levine mismo no piensa que su argumentación pruebe la falsedad del fiscalismo, el fiscalismo que debe tener en mente no es el que tiene pretensiones epistémicas sino uno escindido de éstas. Lo que se muestra, entonces, para Levine, es que el fiscalismo *epistémico*, no el *ontológico* es falso.

Una de las formulaciones del problema de la brecha que este autor hace puede ayudarnos a refinar la distinción que acabo de hacer entre fiscalismo *epistémico* y fiscalismo *ontológico*. Recordemos que para Levine algo era explicado satisfactoriamente cuando ‘no quedaba nada por fuera’ (en el sentido de que se podía

hacer una derivación *completa* de todos los rasgos del *explanandum* a partir del *explanans*). Sin embargo, el mismo Levine considera que esta formulación merece una matización más extensa, relacionada con lo que significa ‘dejar por fuera’; en particular, Levine distingue entre una lectura epistémica y una ontológica de esta expresión.²⁵ Levine dirá que los argumentos de concebibilidad tipo Kripke no muestran que el fisicalismo deje nada por fuera en un sentido ontológico, pero muestran que deja por algo por fuera en un sentido epistemológico. Esto último sucede porque dichos argumentos ponen en evidencia nuestra incapacidad de explicar el carácter cualitativo de las experiencias conscientes en-términos-de/a-partir-de las propiedades físicas de los estados cerebrales.

La *lectura metafísica* de “dejar por fuera” señala que hay objetos, eventos o propiedades a los que no puede referirse la teoría fisicalista. Típicamente, en esta lectura se alega que ciertas situaciones son imaginables, concebibles, etc., y luego se extrae una conclusión ontológica (Levine anota que un primer caso arquetípico de este enfoque sería el de Descartes). En el argumento de Kripke, por ejemplo, se dice que los enunciados de identidad psicofísicos y los que se derivan de reducciones científicas, tienen algo en común: (1) si ambos son ciertos entonces son necesarios y (2) ambos involucran apariencia de contingencia. Pero, según Kripke, hay una importante asimetría entre ambos a la hora de disipar su aparente contingencia: mientras que en un caso se logra descartar esta aparente contingencia, no se logra disipar la imaginabilidad de los casos que falsean los enunciados de identidad psicofísicos (pues la estrategia para descartar la presunta imaginabilidad en los enunciados del primer tipo no funciona con enunciados del segundo tipo). Kripke concluye entonces que no podemos descartar los casos que falsean la identidad psicofísica, lo cual podría constituir evidencia en contra de la tesis misma de que los estados mentales son idénticos a estados físicos.

Sin embargo, Levine piensa que esta conclusión de Kripke se puede evitar sin necesidad de descartar la imaginabilidad de los casos que falsean el enunciado de identidad

²⁵ Podemos anticiparnos de una manera esquemática y señalar que *epistémicamente hablando*, un enfoque ‘deja por fuera’ un fenómeno cuando no se puede hacer una derivación *completa* de todos los rasgos del *explanandum* a partir del *explanans*. *Ontológicamente hablando*, un enfoque ‘deja por fuera’ un fenómeno cuando no llega a referirse a este.

psicofísica, enfatizando que lo que es imaginable es un asunto epistemológico.²⁶ En consecuencia, lo que establece el imaginar –por ejemplo– dolor sin activación de las fibras C es la “posibilidad epistemológica” (i.e., la concebibilidad) de que se dé la una sin que se dé la otra; pero se requiere otro argumento para pasar de la concebibilidad a la posibilidad (“posibilidad metafísica”; que es lo que se requiere para mostrar que no son idénticos).

¿Por qué, entonces, da Kripke por sentado el vínculo entre concebibilidad y posibilidad? Lo que parece yacer aquí, de acuerdo con Levine, es un fuerte sesgo cartesiano: Kripke se estaría orientando por la intuición según la cual cuando se tiene una idea *clara y distinta* se tiene acceso a cómo las cosas son (ontológicamente). De acuerdo con Levine, este parece un modelo verdaderamente optimista de acceso epistémico a la realidad ontológica: no es en absoluto descabellado plantear que nuestras ideas podrían ser tan claras y distintas como quisiéramos, y aún así no corresponder con lo que es de hecho posible (cf. Levine 1993, pp. 123-124). En otras palabras, el mundo está estructurado de cierta manera, lo cual limita sus posibilidades, pero no hay ninguna garantía de que nuestras ideas correspondan apropiadamente con el mundo como es en sí mismo ni que puedan capturar el rango de lo que en él es posible. Si tenemos razón en rechazar la mentada intuición cartesiana, se sigue que la diferencia en la concebibilidad de casos que falsean los dos tipos de enunciados arriba aludidos se vuelve una diferencia irrelevante para el asunto de si dichos casos son o no posibles (y, por ende, para el asunto de si los respectivos enunciados de identidad son idénticos).

Otro ejemplo del tipo de argumentos que entran dentro de lo que Levine llama la ‘lectura metafísica’ es el argumento del conocimiento de Jackson, con espíritu cartesiano semejante al anterior (por lo cual hereda problemas semejantes). Recordemos que Jackson define el fisicalismo como la afirmación de que toda la información correcta es información física. Esta definición, sin embargo, se construye sobre una noción ambigua de ‘información física’, que puede interpretarse como ‘información expresada en términos usados en las ciencias físicas’ o como ‘información sobre hechos físicos’. Lo que el

²⁶ Levine piensa que la apariencia de contingencia radica en que, dado que ninguno de los dos tipos de enunciado es conocido *a priori*, (contra Kripke) **ambos** son *imaginablemente falsos*.

argumento de Jackson demuestra es que la información que Mary adquiere no es física en el primer sentido, pero no demuestra que no lo sea en el segundo sentido. Por lo demás, es sólo en el segundo sentido de información física que el fisicalista está comprometido con la definición de fisicalismo de Jackson: en la medida en que todos los eventos y procesos particulares son eventos y procesos físicos, toda la información correcta es información física.²⁷

Según Levine, el caso de Mary ejemplifica que puede haber más de una manera de tener acceso epistémico a un mismo hecho; y, evidentemente, no se puede pasar siempre directamente de múltiples modos de acceso a múltiples hechos accedidos. Lo que muestra el caso de Mary, de acuerdo con Levine, es que podemos saber (a) qué descripción física satisface un estado mental sin saber (b) cómo es tener tal estado²⁸ (poseer un conocimiento del tipo (a) no garantiza tener un conocimiento del tipo (b)). Pero Levine señala que esto es obvio, pues la única manera de saber cómo es tener un estado mental es teniéndolo; por ende, el 'nuevo conocimiento' de Mary equivale simplemente a su nueva experiencia, que es obviamente nueva en la medida que ella no había tenido esas experiencias dentro de la habitación (y que, como he sugerido, podría llamarse conocimiento en el sentido de 'conocimiento por aprehensión que he explorado páginas atrás).

Levine termina su discusión de la 'lectura metafísica' señalando que lo único que hacen tanto el argumento de Kripke como el de Jackson es poner en evidencia la división epistemológica entre dos modos de acceso a lo que todo lo demás²⁹ indica que es el mismo fenómeno. De un lado, tenemos ciertas descripciones físico-funcionales (extensionales, de tercera persona) de ciertos estados de ciertos sujetos. De otro lado, tenemos descripciones derivadas de la experiencia de primera mano de dichos estados. Pero no tenemos razones para pensar que los argumentos en cuestión muestren que estas dos clases de descripciones no se refieren a los mismos hechos, por lo cual

²⁷ Ya he argumentado por qué creo que ninguna versión del fisicalismo que acepte el primer sentido de información física sería plausible.

²⁸ Es decir, el conocimiento del tipo (a) corresponde a un modo de acceso epistémico diferente e independiente conceptual/epistemológicamente del correspondiente al conocimiento de tipo (b), pero con todo ambos son modos de acceso al mismo hecho.

²⁹ Es decir, múltiples consideraciones causales y metodológicas.

tampoco tenemos razones para concluir que el fisicalismo ‘deja por fuera’ algo en el sentido de que no se refiera de algún modo a tales hechos.

Levine concluye que los argumentos tipo Kripke o Jackson ponen en evidencia nuestra incapacidad de explicar el carácter cualitativo de las experiencias conscientes en términos de las propiedades físicas de los estados cerebrales.³⁰ Podría añadirse que los argumentos de Kripke y Jackson *muestran* que hay modos de darse de un estado que el vocabulario de la ciencia física no logra capturar. Pero no tenemos razones para pensar que dichos argumentos *muestran* que estos modos de darse no corresponden al mismo estado, y que las descripciones correspondientes no se *refieren* al mismo estado; por esto, no tenemos razones para concluir que el fisicalismo ‘deja por fuera’ algo en el sentido de que no se refiera de algún modo a tales hechos.

Ahora bien, uno de los argumentos en defensa de la ‘lectura metafísica’ que Levine examina es aquel que señala que la aparente contingencia de un enunciado de identidad se explica por una distinción real entre dos propiedades del objeto: podemos explicar el estado epistemológico de concebir el mismo objeto como distintos objetos por referencia o dos propiedades distintas mediante las cuales tenemos acceso epistémico a él. He concedido, de la mano de Levine, que para explicar cómo es posible concebir un estado como dos estados distintos, debemos asumir que hay al menos dos ‘modos de presentación’ bajo los que lo aprehendemos: el modo de presentación de ‘primera persona’ y el de ‘tercera persona’. Pero, por el argumento anterior, pareceríamos llevados a afirmar que hay al menos dos propiedades distintas del estado que corresponden a los dos modos de acceso; así, la propiedad por la que identificamos el carácter cualitativo de un estado consciente en el ‘modo de primera persona’ sería diferente de la propiedad por la que identificamos el rol funcional de ese estado en el ‘modo de tercera persona’. De modo que, replica el antifisicalista, las dos clases de propiedades no son idénticas, que es lo que pretende el fisicalismo.

³⁰ Valdría la pena preguntar ¿qué significa explicar el carácter cualitativo de las experiencias conscientes en términos de las propiedades físicas de los estados cerebrales? Sobre esto me extenderé más adelante.

El fiscalista, sin embargo, no necesita recorrer esta línea de argumentación, y puede evitarla al menos de dos maneras diferentes. En la primera (considerada por Levine), se plantea que ciertamente siempre que concebimos un objeto particular de dos maneras diferentes –lo suficientemente distintas como para que creamos que concebimos dos objetos distintos– el objeto en cuestión debe poseer (al menos) dos propiedades distintas que correspondan a estos diferentes modos de presentación. El hecho de que tengamos o no un problema para el fiscalismo depende de cuáles son las dos propiedades distintas con las que nos comprometemos.³¹ En la segunda, como se ha sostenido antes, de la existencia de dos modos de presentación o de acceso no se deriva la existencia de dos entidades o propiedades distintas (así que una y la misma entidad o propiedad puede concebirse bajo dos modos de presentación distintos).

La primera de estas líneas de argumentación, no obstante, puede resultar problemática para el fiscalista. Conceder que al concebir un objeto como dos objetos distintos el objeto debe poseer dos propiedades distintas que correspondan a estos diferentes modos de concebir, conllevaría a admitir que la propiedad de tener un cierto carácter cualitativo no es idéntica a ninguna de las propiedades físicas/funcionales por las que las concebimos el estado bajo el modo de presentación de ‘tercera persona’ (es decir, las capturadas en una explicación neuronal/funcional). Esto significa que la teoría de la identidad sería verdadera de un modo muy extraño: pues el carácter cualitativo podría ser una propiedad física, pero implicaría introducir una clase de propiedades físicas/funcionales que no aparecería en ninguna de las explicaciones neuronales/funcionales actuales o posibles.

2.2.1 La Lectura Epistemológica

De acuerdo con Levine, una teoría fiscalista satisfactoria no debería dar sólo una *descripción* física de los estados y propiedades mentales, sino una *explicación* de ellos; concretamente una explicación de por qué cuando estamos en ciertos estados físicos o

³¹ La idea aquí parece ser que en últimas lo que el fiscalista necesita mantener es que tener cierto carácter cualitativo es una propiedad física. Así, incluso admitiendo que el argumento de arriba pudiera establecer que la propiedad de tener un cierto carácter cualitativo es una propiedad distinta de las propiedades físicas de un estado mental por las que las concebimos bajo el modo de presentación de ‘tercera persona’, no necesariamente tendríamos la clase de conclusión metafísica que persigue el anti-fiscalista (que la propiedad de tener cierto carácter cualitativo no es una propiedad física).

funcionales tenemos las experiencia que tenemos. Esta exigencia, sin embargo, parece excesiva en especial si se tiene en cuenta la pretensión central del fisicalismo tal como lo he caracterizado: postular una tesis ontológica, sobre el estatus de la realidad misma (que se apoyada en ciertas consideraciones particulares, característicamente de tipo causal). Parece claro, según lo hemos visto en la argumentación de Nagel, que el fisicalismo no puede realizar una labor tal como ofrecer una explicación de por qué cuando estamos en ciertos estados físicos o funcionales tenemos las experiencias que tenemos (y, con todo, siguiendo aún a Nagel, esto no significa que el fisicalismo sea falso).

Determinar hasta qué punto una teoría satisfactoria debería depender principalmente de lo que esta teoría pretende hacer y, cuando mucho, de lo que se deriva de esta pretensión central; pero, como he argumentado, no es la pretensión central del fisicalismo –ni se deriva de esta– ofrecer una explicación de por qué cuando estamos en ciertos estados físicos o funcionales tenemos las experiencias que tenemos.³² El requisito que plantea Levine aplicaría para una especie de fisicalismo que tuviera alguna clase de pretensión epistemológica, pero interpretado de manera que sea una tesis ontológica, la necesidad de *explicar* podría no aplicar (cf. Papineau 1998).

Los argumentos de concebibilidad atrás examinados, continúa Levine, demuestran la incapacidad de las teorías fisicalistas para proporcionar una explicación de por qué cuando estamos en ciertos estados físicos o funcionales tenemos precisamente las experiencia que tenemos y no otras. Para Levine, la asimetría entre el caso agua/H₂O y el caso dolor/fibras C radica en que hay una '*necesidad*' que 'fluye' de la reducción del agua al H₂O, que no podemos encontrar en la reducción del dolor a las fibras C. Así pues, Levine afirma que la aparente contingencia del caso dolor/fibras C es bidireccional (es concebible tanto lo primero sin lo segundo como viceversa), mientras que la aparente contingencia de agua/H₂O sólo funciona en el primer sentido: puede ser concebible que el agua no sea H₂O (quizás si tenemos en mente casos tipo Tierra Gemela tal como los analiza Putnam, no Kripke), pero "no es concebible (...) que el H₂O deje de manifestar estas [macropropiedades]" (Levine, 1993, pp. 128). De acuerdo con Levine, la

³² De acuerdo a lo que se ha discutido sobre Nagel, habría un impedimento conceptual-epistemológico de principio para realizar semejante tarea.

divergencia en los dos casos refleja una importante diferencia epistemológica entre las dos reducciones: que la teoría química explica lo que debería explicar, mientras que una teoría fisicalista aún “deja algo por fuera”. En la medida que se deja el carácter cualitativo sin explicar (es decir, se deja sin explicar por qué cuando estamos en ciertos estados físicos o funcionales tenemos las experiencia que tenemos), sigue siendo concebible lo uno sin lo otro.

Ahora bien, recordemos que durante el análisis de Nagel se presentaron las razones por las cuales el fisicalismo ciertamente no puede explicar por qué cuando estamos en ciertos estados físicos o funcionales tenemos las experiencias que tenemos (ni viceversa). Si bien estas razones son difícilmente incontrovertibles (e implican que, en efecto, sigue siendo concebible lo uno sin lo otro), no constituyen un riesgo catastrófico para el fisicalismo (ontológico, por supuesto; véase el apartado sobre Nagel). Por otro lado, en el análisis de la expresión “dejar por fuera” se ha argumentado que el fisicalista puede apelar a cierta interpretación de la expresión para señalar que el fisicalismo (de nuevo, ontológico) no deja por fuera el carácter cualitativo (en un sentido relevante para sus propósitos). Como bien señala Levine, la reducción fenoménico/físico no funciona (como sí lo hace la reducción agua/ H_2O), hay que tener presente que no es la pretensión del fisicalismo llevar a cabo un ejercicio epistemológico de reducción explicativa.

Toda reducción debe explicar lo que es reducido, y la manera en la que podemos decir si esto se ha cumplido es ver si el fenómeno al que se reduce hace epistemológicamente necesario el fenómeno a reducir (si, dados los hechos citados en la reducción, las cosas deben ser de la manera en que aparecen). Ahora, enfatiza Levine, sin importar cuán rica sea la historia del procesamiento de información o neurofisiológico, seguirá resultando completamente coherente imaginar lo uno sin lo otro; pero si la historia física o funcional realmente explicara la experiencia, no sería imaginable.

Identificar el agua con el H_2O permite comprender lo que es el agua al explicar su comportamiento, mientras que en el caso del dolor/activación de las fibras C el carácter

subjetivo de la experiencia queda sin explicar y quedamos con una comprensión *incompleta* del dolor.³³

En términos algo extraños, Levine insiste en que una *reducción* teórica se justifica por su poder explicativo: “*Justificamos la afirmación de que el agua es H₂O al trazar hasta el H₂O la responsabilidad causal y la explicabilidad de las numerosas propiedades superficiales por las que identificamos el agua*” (1993, pp. 131, énfasis añadido). Lo anterior podría interpretarse al menos de dos formas diferentes: (1) que lo que justifica la *reducción* del agua al H₂O es que podamos explicar plenamente las propiedades superficiales por las que identificamos al agua (y rastrear la responsabilidad causal por las mismas) en el H₂O; (2) que lo que justifica la *afirmación* de identidad del agua con el H₂O es que podamos explicar plenamente las propiedades superficiales por las que identificamos al agua (y rastrear la responsabilidad causal por las mismas) en el H₂O.

Según creo, dos cuestiones distintas se están mezclando aquí; para aclarar a qué me refiero, es importante hacer algunas precisiones: cuando hablamos de *reducción* explicativa, lo que queremos decir es que algo (o las propiedades de algo) recibe una explicación completa en virtud de otra cosa (un *explanandum* es explicado en términos de un *explanans*); mientras que cuando formulamos un enunciado de identidad, lo que estamos señalando es que los dos términos de la identidad designan el mismo objeto. En este sentido, Levine acierta al decir que una reducción ‘se justifica’ (o, mejor, *que se ha logrado llevar a cabo una reducción*) cuando logramos explicar plenamente las propiedades de algo por virtud de otra cosa. Sin embargo, cuando estamos tratando con una afirmación de identidad, el asunto de su justificación parece tener algunas peculiaridades que lo diferencian del caso de la justificación de las reducciones explicativas: un enunciado de identidad puede justificarse ‘*parcialmente*’, por decir así, en vista de cierta evidencia empírica o en vista de ciertas consideraciones metodológicas, epistemológicas o conceptuales (por ejemplo, un recurso a la mejor explicación) aunque no se expliquen plenamente los rasgos de uno de los elementos de la identidad en virtud de los rasgos del otro elemento. Esto es, que la afirmación de identidad no sea

³³ Como he señalado anteriormente, esta diferencia es producida por la diferencia correspondiente entre los modos de darse del fenómeno y los entramados conceptuales asociados a cada uno. No se puede esperar que la reducción funcione, como el caso de agua=H₂O (donde no se presenta una diferencia semejante entre los modos de darse del fenómeno en cuestión).

plenamente explicativa no hace que sea injustificada: las razones que la justifican (al menos parcialmente) pueden ser numerosas y variadas (causales, metodológicas, etc., etc.).³⁴

Excursus: El Modelo de reducción de Levine

La reducción del agua al H₂O contiene un obstáculo significativo, anota Levine: que la teoría química y la teoría *folk* no tienen el mismo vocabulario, de manera que en alguna parte será necesario introducir 'leyes puente'. En otras palabras, para obtener una explicación química de por qué el agua hierve o se congela, necesitamos una definición de 'hervir' y 'congelarse' que ponga estos términos en el vocabulario de las teorías a las que se apela en la explicación.

En el enfoque de Levine, una reducción procede en dos etapas: (1) un proceso 'más o menos' *a priori* de 'ajustar' para la reducción el concepto de la propiedad a reducir (identificando el rol causal para el que estamos buscando los mecanismos subyacentes). (2) El trabajo empírico de descubrir cuáles son esos mecanismos subyacentes (cf. Levine 1993, p. 132)

En el caso experiencias conscientes/estados cerebrales, identificar el rol causal de las primeras no parece tener mayor dificultad, por lo que parecería que se ha llevado a cabo la primera etapa de la reducción. Lo mismo podría decirse para la segunda etapa. Si lo único que hubiera en nuestro concepto de Experiencia Consciente fuera su rol causal, entonces la labor estaría completada, y no habría diferencia entre la reducción teórica del agua y la de las experiencias conscientes. Pero parece haber más en el concepto de una experiencia consciente (por ejemplo, el dolor) que su rol causal, está su carácter cualitativo o fenoménico: el hecho de que podamos imaginarnos el uno sin el otro muestra que el rol causal no es todo lo que hay en nuestro concepto de experiencia consciente.

³⁴ Recordemos las afirmaciones de Nagel en el sentido de que podemos tener evidencia acerca de una identidad que no entendemos: las identidades no tienen que ser plenamente explicativas, basta con que existan consideraciones sólidas que las respalden.

En este sentido, no puede hacerse una reducción de la experiencia consciente en términos de su rol funcional; y esto parece deberse al hecho de que nuestros conceptos de experiencias conscientes no representan única o principalmente roles causales.³⁵

2.3 CHALMERS: EL PROBLEMA DIFÍCIL, LA IMPOSIBILIDAD DE LA REDUCCIÓN Y EL PANPSIQUISMO

2.3.1 EL PROBLEMA DIFÍCIL Y LA IMPOSIBILIDAD DE LA REDUCCIÓN

Las afirmaciones con las que Chalmers (1995, 1996) empieza podrían resultarle al lector desprevenido lugares comunes en la discusión filosófica sobre la experiencia consciente. Por un lado, insiste en que los métodos reductivos inevitablemente fracasan al tratar de hacerle frente al problema de la conciencia, por lo cual se requiere una nueva clase de explicación no-reductiva. También hace énfasis en el punto –en el que paradójicamente converge con Paul Churchland– según el cual el término “conciencia” es muy ambiguo y se utiliza para designar muchos fenómenos diferentes. No pasa, sin embargo (como sí lo hace Churchland), a decir que un término de estas características debería ser eliminado; más bien plantea la necesidad de explicar cada uno de los fenómenos que caen bajo el término “conciencia”. Es allí donde aparece una de las tesis centrales de Chalmers, que de los fenómenos en cuestión, algunos son más fáciles de explicar que otros. Los fáciles son los que parecen abordables mediante los métodos de explicación estándar de la ciencia, en los que los fenómenos son explicados en términos de mecanismos neuronales o computacionales (es decir como el cumplimiento de funciones); muy distintos son los difíciles (o mejor, *el* difícil), pues parecen resistir dichos métodos.

Es la subjetividad de la experiencia consciente la que produce lo que Chalmers considera el *Problema Difícil* de la conciencia: cuando pensamos y percibimos, hay diversas formas de procesamiento de información, pero también hay un aspecto subjetivo (hay una

³⁵ Además, de la introducción de casos tipo Tierra–Gemela, Block y Stalnaker (1998) ha controvertido ampliamente la noción de reducción de Levine argumentando que la etapa (1) no es más o menos *a priori* (sino que no lo es de modo alguno), lo cual haría defectuoso el modelo de reducción de Levine.

manera como se siente ser un organismo consciente; cf. Nagel 1974). Surgen entonces una serie de preguntas como ¿por qué cierto procesamiento físico (o la realización de cierta función) debería estar acompañado por experiencias conscientes? El Problema Difícil es responder esta pregunta (en adelante, pregunta C).

Si bien es cierto que todo lo que podría involucrar la explicación de un *problema fácil* es en principio una explicación de cómo se realiza la función relevante, una explicación de este tipo no responde a la pregunta C. Esto se muestra en el hecho de que incluso cuando hemos explicado la realización de todas las funciones cognitivas relevantes, aún podemos hacer la pregunta C (esta pregunta no se plantea en ninguno de los problemas fáciles, y plantearla se vería más bien como un error conceptual). De una explicación neurológica o funcional sólo se derivarán más funciones o procesos neurológicos, así que los hechos de la experiencia no pueden ser una consecuencia de ninguna explicación de esta clase. Por consiguiente, el *problema difícil* no es un problema acerca de la realización de funciones; lo esencial de la subjetividad de la experiencia consciente, o conciencia fenoménica, no es caracterizable como el cumplimiento de una función. Las caracterizaciones físicas o funcionales de la conciencia no dicen ni pueden decir nada acerca del carácter subjetivo de un estado mental: la descripción más exhaustiva del cumplimiento de una función, no habrá dicho nada sobre el aspecto subjetivo. Dicho de otro modo, ninguna explicación de la conciencia dada en términos de la ciencia física puede responder la pregunta “¿por qué este proceso debería dar lugar a la conciencia?” pues siempre es conceptualmente coherente que pudiera darse sin experiencia consciente. La subjetividad de la experiencia consciente va más allá de lo que puede derivarse a partir de la ciencia física.

Chalmers se refiere a varios trabajos de investigación sobre la conciencia, observando que *asumen* la conexión entre el cerebro y la experiencia, en lugar de *explicarlo*. De acuerdo con Chalmers, necesitamos poder deducir por qué debería haber experiencia: “una teoría completa de la conciencia debe edificar un puente explicativo” (Chalmers, 1995, p.207). Es claro que la explicación de funciones no basta para la explicación de la experiencia, y este es un punto conceptual básico que ninguna investigación neurocientífica puede afectar.

2.3.2 La Dificultad de Entender el Problema Difícil

La dificultad al momento de interpretar el Problema Difícil de Chalmers, parece derivarse de las múltiples formulaciones que hace del problema. Estas formulaciones no son simples parafraseos, sino que parecieran hacer énfasis distintos y, por tanto, parecieran estar haciendo preguntas distintas.

Por ejemplo, cuando pregunta “¿Cómo es que estos sistemas [físicos] son sujetos de experiencia?” (p.202) pareciera estar haciendo una pregunta de la forma ¿cómo es que la entidad x del mundo tiene la propiedad P ? A preguntas de este tipo parece difícil encontrarles respuestas que no parezcan una perogrullada. Si preguntamos ¿cómo es que la luz tiene la propiedad de viajar a una velocidad de trescientos mil kilómetros por segundo? ¿Por qué la constante de Plank (h) es igual a $6,626 \times 10^{-34}$ julios por segundo? Parece que a preguntas como esta no podríamos responder sino con la afirmación de que así es el mundo, simplemente la luz tiene tal velocidad y la constante de Plank tiene tal valor. Algo semejante se puede decir de otras formulaciones: “¿Por qué es que cuando nuestros sistemas cognitivos entran en procesamiento de información visual y auditivo, tenemos experiencia visual o auditiva (...)?” (p.202), “¿cómo podemos explicar por qué hay algo que es considerar una imagen mental, o experimentar una emoción?” (Ibíd.), “¿por qué y cómo surge la experiencia de una base física?” (Ibíd.), “¿por qué la ejecución de estas funciones es acompañada por la experiencia?” (p.204).

Esto ha motivado a teóricos como David Papineau a señalar que este tipo de preguntas son semejantes a la pregunta “¿Por qué son idénticos a y b ?”, la cual no cabe hacer. Si $a=b$, entonces simplemente lo son, no cabe preguntar por qué son idénticos.

Pienso que sólo una de las formulaciones que elige Chalmers muestra la brecha de una forma más o menos clara y es menos vulnerable a lecturas como la que hace Papineau: “¿Por qué el procesamiento físico debería originar una rica vida interna?” (Chalmers, 1995, p.202). Si no hubiera brecha explicativa, la descripción del procesamiento cerebral en términos neurológicos debería decirnos por qué tiene tal propiedad fenoménica (de la misma manera en que la descripción del H_2O en términos de la ciencia física nos dice por qué el agua tiene tal propiedad); pero esto es precisamente lo que no sucede, queda sin explicar por qué el procesamiento cerebral tiene tal propiedad fenoménica. Como he señalado en mi discusión de Levine, la brecha consiste en que algo es dejado por fuera

(y *no puede* ser incluido) en el sentido de que no se pueda hacer una derivación de los hechos de la experiencia a partir de hechos neurofisiológicos. He señalado también que puede que nunca lleguemos a entenderlo: en una línea semejante a la de Collin McGinn (y Thomas Nagel), puede que nuestras limitaciones cognitivas–epistemológicas sean tales que no podamos entender que el fisicalismo sea correcto, pero esto no significa que no sea así. Puede que la brecha nunca se cierre, pero ese no es un problema para el fisicalista; mi sugerencia ha sido que la respuesta del fisicalista podría ser la siguiente: los hechos de la experiencia no se pueden derivar de hechos neurofisiológicos no porque no sean ontológicamente idénticos, sino porque corresponden cada uno a un esquema, marco conceptual o modo de concebir diferente. El fisicalista debería estar de acuerdo con toda esta intuición *epistemológica*, de una explicación que no es capaz de capturar un fenómeno.

La fuerza de esta formulación de la pregunta también parece derivarse de la siguiente consideración: que no es razonable decir que dichos procesos *deberían* dar lugar a la conciencia fenoménica; Chalmers basa la mencionada consideración diciendo que bien podríamos imaginar que esta función se llevara a cabo *en la oscuridad* (“*Dado cualquier proceso tal, es conceptualmente coherente que pudiera instanciarse en la ausencia de experiencia*” p. 209). Ya he discutido algunas razones que problematizan este tipo de argumentos de concebibilidad, en especial la idea de que la relación coherencia conceptual y posibilidad metafísica es tremendamente oscura y, quizás, no sea de implicación.³⁶

2.3.3 El Problema Difícil y la Brecha Explicativa

Algo que cabe suponer es que el problema *de principio* de que no se pueda hacer una derivación de los hechos subjetivos de la experiencia a partir de hechos neurofisiológicos constituye el principal impedimento para responder la pregunta C. En otras palabras, lo que hace difícil el *problema difícil* es que hay una brecha explicativa. Como he

³⁶ También se ha alegado (Dennett, 1991) que cuando se habla concebibilidad en muchos casos se trata simplemente una serie de actos de ‘súper-imaginación’ (actos en los que realmente el supuesto concipiente no logra imaginar nada aunque manifieste haberlo hecho). Este argumento, a mi juicio, contiene una serie de defectos por los cuales no cumple su cometido, pero no es este el lugar para extenderse en su estudio.

argumentado en mi discusión del enfoque de Levine, creo es que esta es una brecha conceptualmente insalvable, de naturaleza epistémológica más que ontológica, que se da entre las funciones o lo físico y la experiencia (producida por la independencia lógica entre dos ‘modos de concebir’ o modos de darse de –lo que parece ser– un mismo fenómeno); de ser así, la pregunta C no se puede responder. Y si fuera verdad lo que dice Chalmers con respecto a que “*una teoría completa de la conciencia debe edificar un puente explicativo*” entre las funciones o lo físico y la experiencia, entonces no puede existir una teoría satisfactoria de la conciencia.

Algunas consideraciones a propósito del tipo de conocimiento que se requiere para cerrar la brecha explicativa, apoyarían lo que acabo de decir sobre su insalvabilidad. Pensamos que no hay una brecha explicativa entre el agua y el H₂O porque podemos derivar las propiedades del agua a partir de las propiedades del H₂O, es decir, porque podemos conocer las propiedades de uno de los idénticos a partir de las propiedades del otro.³⁷ Pongamos ahora lo anterior en términos de la brecha explicativa entre lo cerebral y la experiencia consciente, para cerrar esta brecha se requeriría que pudiéramos conocer las propiedades de uno de los idénticos a partir de las propiedades del otro, es decir, que pudiéramos conocer las propiedades de la experiencia consciente (de dolor, por ejemplo) a partir de las propiedades de cierta actividad cerebral. ¿Qué es lo que se requiere entonces para cerrar la brecha explicativa? Que, por ejemplo, a partir del conocimiento de las propiedades de cierta actividad cerebral, alguien tenga conocimiento de cómo se siente el dolor (incluso sin haber sentido nunca dolor). ¡Pero el conocimiento de cómo se siente el dolor es el conocimiento subjetivo, de primera persona, una sensación! Así que el tipo de conocimiento que se requiere para cerrar la brecha es tal que le debería permitir saber a alguien como qué se siente el dolor (incluso aunque nunca haya sentido dolor). Debe ser un conocimiento de neurología dado desde la perspectiva de tercera persona, que me permita conocer una sensación desde la perspectiva de primera persona; es como si al adquirir ese conocimiento de neurología, sintiera dolor. Si no sintiera el dolor ¿qué significa decir que tengo un conocimiento que me permite saber cómo es tener una experiencia que no he tenido? Recordemos a Jackson, que cuando

³⁷ Alguien que, por ejemplo, no conozca las propiedades del agua podría conocerlas a partir del conocimiento de las propiedades del H₂O

termina de discutir el argumento de Nagel señala que si el fisicalismo fuera verdadero, suficiente información física acerca de Fred obviaría cualquier necesidad de extrapolar o realizar actos especiales de imaginación o entendimiento para saber todo acerca de su experiencia especial de color.

Lo absurdo de hablar de un conocimiento de neurología que me permitiera sentir dolor es que este conocimiento no puede existir. Si es verdad que cerrar la brecha requeriría una aproximación que nos permitiera pasar del modo de concebir 'de primer persona' al modo de concebir 'de tercera persona', sucede que este conocimiento involucraría exigencias contradictorias: suprimir el punto de vista del sujeto y, al mismo tiempo, incluirlo. No se puede establecer un flujo explicativo de este tipo, desde (en palabras de Nagel) el modo de concebir subjetivo al modo de concebir objetivo; no se puede incluir el modo de concebir de la primera persona en el de tercera persona, ni se puede pasar súbitamente del uno al otro sin dar problemáticos saltos. Cuando mucho, repetiríamos lo que Chalmers llama hacer que "la experiencia repentinamente entre en la imagen" (Chalmers, 1995, p.207). Un *tránsito explicativo* del modo de concebir de primera persona al modo de concebir de tercera persona (o viceversa), requeriría que sacrificáramos algo que es definitorio de cada una de estas maneras de concebir (en un caso, el punto de vista del sujeto, en el otro caso la eliminación del mismo). Se sigue no sólo que existe la brecha explicativa, sino que es insalvable.

Las observaciones que he hecho durante toda la revisión de la posición de Chalmers apuntan precisamente a la incapacidad de su teoría para hacer lo que él considera que *tiene* que hacer una teoría aceptable sobre la conciencia; esto sucede porque, de hecho, *ninguna* teoría sobre la conciencia puede hacer lo que Chalmers cree que tiene que hacer.

Siguiendo a Nagel, esta incapacidad para cerrar la brecha no significa que el fisicalismo sea falso. Es verdad que una explicación funcional o de la ciencia física sólo da funciones, y que los hechos de la experiencia no pueden ser consecuencia de ninguna explicación tal. Pero este es un problema conceptual/epistemológico, de marcos conceptuales; esto es lo que hace que sea *conceptualmente/epistemológicamente coherente* que cualquier proceso funcional pueda darse sin experiencia (pues los dos tipos de términos hacen parte de esquemas conceptualmente independientes). En efecto,

la experiencia puede surgir de lo físico, pero no se *deriva* de lo capturado por las ciencias físicas.

Así, una teoría de la experiencia consciente bien podría ser aquella que explique por qué la brecha existe y por qué es insalvable, mantiene, si es el caso, una ontología monista fisicalista. Si bien el fisicalista puede tener razones para creer que lo físico y lo fenoménico son dos modos de darse de una misma realidad, y puede apelar a razones de tipo causal para sostener que así es, puede que nunca lleguemos a entender cómo es posible que sea el caso. Hay una brecha explicativa insalvable, probablemente por el alcance de nuestra estructura cognitiva (McGinn), porque se trata de dos modos de concebir con sistemas conceptuales asociados independientes entre sí.

No hay duda de que, en cierto modo, Chalmers está en lo correcto cuando señala que “para explicar la experiencia consciente, necesitamos un ingrediente extra en la explicación” (Chalmers, 1995, p.208). Sin embargo, ninguna de las propuestas de ingrediente *extra* permite el tránsito explicativo una concepción subjetiva, fenomenológica, que es lo que se requiere para que no quede nada *por fuera*. Lo que estoy proponiendo, entonces, es que este ingrediente extra tiene un carácter epistemológico; no se trata de un dominio ontológico *extra*, no sólo porque sea innecesario (puede prescindirse de él), sino porque su introducción no resuelve los problemas con los que empezamos.

Por otro lado, la relación del Problema Difícil con la brecha explicativa permite conjeturar que el problema de Chalmers es de corte conceptual/epistemológico, algo que no se resuelve con una nueva ontología. ¿Por qué el hecho de que los métodos reductivos sean impotentes para explicar algo que no es una función debe tener consecuencias ontológicas y no simplemente epistemológicas? La teoría de Chalmers no sólo sería ontológicamente innecesaria sino epistemológicamente insuficiente. He planteado varias dudas sobre si la propuesta de Chalmers es capaz de responder la pregunta que origina el *Problema Difícil* y es capaz de construir semejante *puente explicativo*. Si su propuesta fuera en efecto incapaz de resolver el problema, o si la exigencia de hacerlo dejara de valer en algún momento, entonces sus críticas a otras aproximaciones perderían vigor o se aplicarían a su propia teoría.

2.3.4 LA TEORÍA PANPSIQUISTA DE CHALMERS

Llama la atención el hecho de que Chalmers ponga restricciones tan severas para la plausibilidad de las teorías de la conciencia, como esta (refiriéndose al enfoque de Crick y Koch):

Cuando mucho, da por sentada la existencia de la experiencia y explica algunos hechos acerca de su estructura, proporcionando un tipo de explicación no-reductiva de los aspectos estructurales de la experiencia (...) Esto es útil para muchos propósitos, pero no nos dice nada acerca de por qué debería haber experiencia (...).(Chalmers, 1995, p.207)

Como podemos ver, la exigencia –o al menos una exigencia central– de Chalmers para que una teoría de la conciencia fuera satisfactoria era que nos ofreciera una respuesta a la pregunta C. Pero ahora parece que su teoría puede ser admisible aunque renuncie a responder dicha pregunta:

Por supuesto, al tomar la experiencia como fundamental, hay un sentido en el que esta aproximación no nos dice por qué hay experiencia consciente (...). (Chalmers, 1995, p.211)

Así, mientras que ofrecer una respuesta completa y cabal para la pregunta C era una *conditio sine qua non* para todas las demás teorías de la conciencia –si pretendían llegar a ser teorías *satisfactorias* de la conciencia–, no lo es para la aproximación del propio Chalmers.

Recordemos que Chalmers admite que hay una brecha explicativa '*entre las funciones y la experiencia*', y afirma que necesitamos un *punte explicativo* para salvar la brecha. En este sentido, indica que está considerando la respuesta a la pregunta C como la que sirve de puente explicativo entre lo físico y lo fenoménico. Con todo, parece haber serias dudas sobre lo que Chalmers pretende hacer con la brecha explicativa: Al tomar la experiencia como fundamental esta aproximación no nos dice por qué debería haber experiencia ¿cómo tiende entonces el puente? ¿Cómo puede ser explicativo convertir el explanandum fundamental (es decir, eliminando por definición la posibilidad de un *explanans*)?

Veamos cómo justifica este movimiento. Como se anticipa en la cita inmediatamente anterior, Chalmers postula que una teoría de la conciencia debería tomar la experiencia *misma* como una propiedad fundamental de la realidad. Si todo en la teoría física es

compatible con la ausencia de conciencia, entonces una teoría de la conciencia requiere la adición de algo ontológicamente fundamental.

Con todo, no queda muy claro lo que significa tomar *la experiencia misma* como una propiedad fundamental de la realidad (tal como la masa o la carga eléctrica); parece significar algo en estas líneas: que hay propiedades fenoménicas fundamentales de nivel micro (llamémoslas ‘micro–fenoménicas’, y serían el equivalente de las propiedades físicas fundamentales de nivel micro, o micro–físicas), y que estas propiedades ‘micro–fenoménicas’ (de manera análoga a como las propiedades micro–físicas dan lugar a propiedades macro–físicas) podrían dar lugar –bajo condiciones adecuadas de combinación– a una conciencia fenoménica como la nuestra. Pero la posición de Chalmers enfrenta aquí el siguiente dilema: o bien las propiedades micro–fenoménicas tienen el mismo tipo de sensación cualitativa característica de la conciencia fenoménica ordinaria, o no lo tienen. Si lo tienen, parece increíblemente difícil entender cómo podría una de estas propiedades micro–fenoménicas tener una sensación cualitativa ordinaria (vg, un dolor). Y si las propiedades micro–fenoménicas no tienen tales sensaciones cualitativas, no es claro por qué deberían ser mejores candidatas que las propiedades físicas para explicar las experiencias conscientes ordinarias.³⁸ Lo que sí parece quedar claro de las afirmaciones de Chalmers es que se estaría considerando lo fenoménico como otro dominio ontológico, no–físico (resulta difícil entonces dar sentido a las afirmaciones de Chalmers sobre la simplicidad y elegancia de esta posición, en especial en la medida que es una violación la navaja de Ockham).

³⁸ ¿Cómo pasamos explicativamente de estos *quanta* de conciencia a la conciencia que nos es familiar? ¿Hay alguna contradicción en la idea de que estas propiedades micro–fenoménicas podrían presentarse en la ausencia de experiencia consciente a nivel macroscópico? ¿Cómo evitar que se repita la brecha explicativa?

Para evitar este problema parecería que debe tomar *nuestras experiencias* como fundamentales, no ciertos *quanta* de experiencia.

Además, la respuesta que da Chalmers a la pregunta C describe una serie de interacciones entre elementos básicos (“micro–fenoménicos”), interacciones que dan como resultado la capacidad de tener experiencias conscientes; se sigue entonces que las interacciones que preceden esta capacidad de tener experiencias conscientes deben describirse desde el punto de vista de tercera persona (pues preceden a la existencia del sujeto consciente en cuestión), pero las experiencias conscientes se describen desde el punto de vista de primera persona.³⁸ Así que si pasáramos del nivel de las interacciones entre elementos básicos al nivel de las experiencias conscientes de ciertos sujetos, tendríamos un caso en el que la experiencia de primera persona aparece *de repente*.

Ahora bien, muchas de las consideraciones que ofrece Chalmers al presentar su teoría resultan problemáticas y confusas. Observa que toda propiedad fundamental se acompaña de ciertas leyes fundamentales que dicen, por ejemplo, cómo se relaciona la experiencia con los rasgos físicos del mundo. Estos “principios psicofísicos” deberían decirnos qué clase de sistemas físicos tendrán asociadas experiencias (y en el caso de sistemas que las tienen, deberían decirnos qué clase de propiedades físicas son relevantes para la aparición de experiencia, y qué clase de experiencia deberíamos esperar que sea proporcionada por un sistema físico dado).

El primero de ellos es el *Principio de Coherencia Estructural*, el cual señala que hay una coherencia entre la estructura de la conciencia y la estructura del procesamiento de información. Generalmente donde encontramos la una, encontramos la otra. La experiencia tiene una estructura compleja: para cada uno de estos rasgos estructurales, hay un rasgo correspondiente en la estructura de procesamiento de información de la experiencia.

Pero este principio, señala Chalmers, tiene sus límites (unos que, de por sí, parecen restringir bastante el alcance del principio): aunque nos permite recuperar las propiedades estructurales de la experiencia a partir de propiedades del procesamiento de información, no todas las propiedades de la experiencia son propiedades estructurales; algunas propiedades de la experiencia (como la naturaleza intrínseca de una sensación de rojo) no pueden ser capturadas completamente en una descripción estructural. Esto significa que las propiedades estructurales de la experiencia restringen la experiencia sin agotarla.³⁹

El segundo es el *Principio de Invariancia Organizacional*, el cual señala que cualesquiera dos sistemas con la misma organización funcional detallada tendrán experiencias cualitativamente idénticas.

No queda muy claro cuál es el estatus y de los principios psicofísicos que propone Chalmers, ni cuál es su papel. Claramente no nos permiten entender que cierta clase de

³⁹ A decir verdad, el caso de los qualia invertidos muestra que las propiedades estructurales no tienen un papel muy restrictivo realmente.

sistemas físicos tengan asociadas experiencias, ni entender que cierta clase de propiedades físicas sean relevantes para la aparición de experiencia, ni entender que deberíamos esperar que cierta clase de experiencia sea proporcionada por un sistema físico dado. Lo anterior resulta de algo que Chalmers anota a propósito de uno de estos principios: *“no es un principio lógicamente necesario, puesto que después de todo podemos imaginar que todo el procesamiento de información ocurre sin experiencia alguna, pero sin embargo es una fuerte y familiar restricción sobre la conexión psicofísica”* (Chalmers, 1995, p.214) y que *“refleja el hecho central de que incluso aunque de los procesos cognitivos no se deriven hechos acerca de la experiencia consciente, la conciencia y la cognición no flotan libres una de la otra sino que van juntas de una manera íntima”* (Ibíd). Describen entonces una serie de ‘armonías preestablecidas’ y no tienen un carácter explicativo. Podríamos preguntar también ¿por qué considerarlos leyes que acompañan una propiedad fundamental si no se cumplen necesariamente, sino sólo en un rango de casos muy restringido (lo que llamamos los casos “más familiares”)?

Además ¿cómo podrían estos principios ayudar a cerrar la brecha? Recordemos que el tipo de conocimiento que se requiere para cerrar la brecha debería hacer inteligible el hecho de que de cierta estructura física se derive tal o cual experiencia (y esto, parece, no se puede hacer). Se requiere algo que nos permita hacer un tránsito natural desde la perspectiva de tercera persona a la de primera persona. Para cerrar la brecha es insuficiente la descripción de un conjunto de meras correlaciones; las correlaciones entre un lado de la brecha y otro lado de la brecha no tienen un carácter *explicativo*.

La noción de correlación entre propiedades estructurales de la experiencia y propiedades estructurales del procesamiento de información da lugar a una sospecha más. El tipo de conocimiento que se requiere para cerrar la brecha debería permitirme entender la experiencia en cuanto tal, totalmente. Una comprensión de manera esquemática, puramente correlacional y estructural también la podría proporcionar la neurociencia.

Una vez presentados los principios psicofísicos, Chalmers expone su ‘*Teoría del doble aspecto de la información*’, donde afirma que la información tiene dos aspectos básicos: uno físico y otro fenoménico. En la medida que la información es fundamental para la física del universo, se llega a una concepción del mundo en la que la información es fundamental y en la que tiene dos aspectos básicos: los rasgos físicos y los fenoménicos del mundo.

Chalmers también admite, además, una versión del panpsiquismo, pues ante la pregunta de si toda la información tiene un aspecto fenoménico, Chalmers reconoce que a esta respuesta se puede responder negativamente pero piensa que es igualmente interesante y plausible una respuesta afirmativa (que la experiencia consciente esté tan extendida como la información). Si toda la información tiene un aspecto fenoménico y la información subyace a todo lo físico (a todas las entidades físicas) entonces todo tendría experiencias conscientes. Chalmers admite que incluso los termostatos podrían tener experiencias, pero lo que se seguiría es que todas las entidades físicas podrían tenerlas. ¿Qué impediría que hubiera experiencias conscientes de moléculas, de glóbulos rojos, de campos electromagnéticos, etc.? ¿Cómo es ser un campo electromagnético?

Chalmers indica además que los principios psicofísicos no interferirán con las leyes físicas, en la medida que parece que las leyes físicas describen un sistema cerrado causalmente (afirmando además explícitamente que esta es una versión del dualismo compatible con lo demás que sabemos del mundo).

Estas afirmaciones, sin embargo, deben ser revisadas con precaución. Que estos principios *psicofísicos* no interfieran con leyes físicas, puede significar dos cosas: que las 'leyes psicofísicas' describen puras correlaciones entre dos dominios ontológicos que no se afectan causalmente (armonías preestablecidas sin ningún carácter explicativo), o bien que describen relaciones causales entre dos dominios ontológicos, lo cual interfiere –viola– la ley del cerramiento causal del mundo físico. Así que si estos principios psicofísicos son un simple addendum a nuestra imagen científica del mundo, no tienen mayor plausibilidad que la doctrina del paralelismo psicofísico, y si no son un simple addendum entonces interfieren con las leyes físicas, son incompatibles con la física y con la visión científica del mundo.

En otras palabras, incluso a primera vista este enfoque parece contradecir leyes básicas de la física como la del cerramiento causal del mundo físico: según esta ley, lo físico sólo interactúa causalmente con lo que está dentro de su mismo dominio ontológico, así que o bien lo fenoménico no tiene poder causal (como en el paralelismo y el epifenomenalismo), o bien sí lo tiene (por lo cual tendría que ser físico o tendría que violarse el principio).

No hay duda de que la propuesta de Chalmers es una forma de dualismo, pero tampoco debería haber duda de que no es en absoluto un dualismo “*inocente (...) enteramente compatible con la visión científica del mundo*” ni de que es falso que “*nada en esta aproximación contradice nada en la teoría física*” (Chalmers, 1995, p.211).

2.3.5 La experiencia consciente como propiedad fundamental de la realidad

Chalmers señala que los físicos se han visto llevados muchas veces a introducir propiedades fundamentales. Sin embargo, los físicos citan la evidencia que los ha conducido a introducir estas categorías como fundamentales; pero lo único que encontramos de este talante en Chalmers es la idea de que en la medida que la conciencia fenoménica no puede ser reducida *explicativamente*, debe ser una propiedad fundamental.

Podemos ver cuán débilmente está planteada la experiencia como propiedad fundamental mediante la siguiente comparación: supongamos que alguien propone que dado que algunas cosas son simplemente lindas y otras cosas no lo son, y dado que todos podemos notarlo (aunque resulta muy difícil describirlo e incluso parece imposible explicarlo mediante métodos funcionales), deberíamos postular la ‘lindura’ como una propiedad fundamental de la realidad (así como la masa, la carga eléctrica, etc.). La lindura claramente no es una propiedad funcional: puedo imaginar que alguien que no es lindo se comporte como si fuera lindo (no es un caso tan imaginario después de todo). ¿Que no podamos explicar la lindura con los métodos tradicionales es señal de que es ontológicamente fundamental? Parece que la lindura involucra una forma de ver, una perspectiva, no una propiedad fundamental.

Podríamos caracterizar, más bien, los aspectos ‘fáciles’ de la conciencia simplemente como aquellos que podemos capturar y describir desde una perspectiva de tercera persona (i.e., física o funcionalmente), mientras que la experiencia consciente (el aspecto ‘difícil’ de la conciencia) simplemente sería aquel que podemos capturar y describir únicamente desde el punto de vista de la primera persona. Mediante los métodos tradicionales de tercera persona (i.e., los físicos o funcionales) podríamos describir todos

los aspectos capturables desde la tercera persona, pero no podríamos capturar nada esencial acerca del aspecto capturable únicamente desde la primera persona (la experiencia consciente, entonces, sólo sería capturable con métodos de primera persona como la fenomenología). Ahora, caracterizados de esta manera los dos aspectos (el de primera y el de tercera persona) bien podrían ser aspectos o perspectivas del mismo fenómeno (tendríamos un dualismo epistemológico que conserva el monismo ontológico).

Así, no sería necesario ampliar el mobiliario del universo, pues se requeriría una nueva aproximación epistemológica y no ontológica del problema. Es más existen varias afirmaciones de Chalmers que nos hacen pensar que podría bastar con quedarnos en el nivel epistemológico y que no hay ninguna necesidad de pasar al plano ontológico:

Desafortunadamente, hay razones sistemáticas por las que estos métodos deben fallar. Los métodos reductivos son exitosos en la mayoría de dominios porque lo que necesita explicarse en esos dominios son estructuras y funciones, y estas son la clase de cosa que una explicación física puede establecer. Cuando llega a un problema por encima de la explicación de estructuras y funciones, estos métodos son impotentes. (Chalmers, 1995, p.209)

La clave (...) es el punto conceptual de que la explicación de funciones no basta para la explicación de la experiencia. Este punto conceptual básico no es algo que la investigación neurocientífica posterior afectará. (Chalmers, 1995, p.209)

Más aún, cuando Chalmers trata de hacer afirmaciones ontológicas basándose en las anteriormente citadas consideraciones epistemológicas, da la impresión de un *non-sequitur*.

Sabemos (sic) que una teoría de la conciencia requiere la adición de algo fundamental a nuestra ontología, en la medida que todo en la teoría física es compatible con la ausencia de conciencia. (Chalmers, 1995, p.210)

La admisión del pansiquismo o la postulación de la información como una propiedad fundamental de la realidad, implican cambios absolutamente drásticos y radicales en nuestra concepción del mundo. Con todo, lo que he querido esbozar es que parece que no necesitamos utilizar esta estrategia y que podrían existir maneras de entender la conciencia que no impliquen derrumbar nuestra imagen del mundo.

2.4 LOAR Y LOS CONCEPTOS FENOMÉNICOS

Ante la afirmación de una brecha explicativa entre los procesos físicos y las experiencias conscientes, las reacciones de los filósofos han sido muchas y muy variadas: (a) negar que exista la brecha y sostener que se trata de cierto tipo de ilusión; (b) afirmar que hay una brecha que en el futuro se cerrará; (c) sostener que la brecha explicativa es síntoma de una brecha ontológica; (d) señalar que la brecha es un mero resultado de nuestra manera de concebir la conciencia. Lo interesante de esta última posición no localiza la brecha en la relación que existe entre los procesos físicos y las experiencias conscientes, sino en la relación que existe entre nuestros conceptos de los procesos físicos y nuestros conceptos de las experiencias conscientes.

Los defensores de esta estrategia argumentan que los conceptos fenoménicos (nuestros conceptos de las cualidades fenoménicas) tienen una naturaleza especial. Sugieren que dada esta naturaleza especial, es predecible que encontremos una brecha explicativa entre los procesos físicos –concebidos bajo conceptos físicos– y las experiencias conscientes –concebidas bajo conceptos fenoménicos–. De ser correcta, esta posición implicaría que podemos no tener una explicación física de la conciencia fenoménica, de la experiencia consciente (es decir, puede que nunca podamos cerrar la brecha explicativa), pero que podemos seguir siendo metafísicamente fisicalistas (contando, además, con una explicación de por qué la existencia de la brecha explicativa no afectaría la verdad del fisicalismo).

Adicionalmente, esta estrategia estima ciertas intuiciones que han conducido a algunos a rechazar el fisicalismo como consecuencia de la naturaleza especial de los conceptos fenoménicos, por lo que no pueden constituir una razón concluyente para rechazar el fisicalismo. Esta reciente opción parece altamente atractiva para el fisicalista, pues respeta tanto la realidad de la experiencia consciente con todas sus características subjetivas como la tesis metafísica del fisicalismo, explicando además las intuiciones epistemológicas que generan el problema de la conciencia fenoménica (y mostrando que tales intuiciones no riñen con el fisicalismo).

Brian Loar, uno de los principales proponente de esta estrategia, desarrolla buena parte de su enfoque en el artículo “Phenomenal States” de 1990 y en “Phenomenal States:

Second Version” de 1997. En lo que sigue, presentaré la estrategia de los conceptos fenoménicos basándome en la segunda versión de la posición de Loar. Procuraré dejar sentados lo más claramente posible los conceptos y los argumentos que presenta, con miras a evaluar en el éxito o fracaso de la estrategia de los conceptos fenoménicos.

2.4.1 CONCEPTOS Y CUALIDADES: LO FÍSICO Y LO FENOMÉNICO.

Todos alguna vez hemos sentido el pinchazo de una aguja, y nos parece claramente diferenciable de un dolor de espalda, y este, a su vez, de un dolor de muelas. Alguna vez hemos sentido cosquillas, las cuales podemos diferenciar fácilmente de la comezón; hemos sentido euforia, o hemos estado tristes. ¿En qué se diferencian todas estas experiencias conscientes? Según Loar, se diferencian en *cómo se sienten*. Es así como Loar empieza su una caracterización preliminar de las nociones de *cualidad fenoménica* y *concepto fenoménico*. Las *cualidades fenoménicas* son las cualidades de *nuestras experiencias conscientes*, mediante las cuales diferenciamos nuestras experiencias por *cómo se siente tenerlas*; las concepciones que tenemos de estas cualidades son los *conceptos fenoménicos*. Estos conceptos son formados a partir del propio caso, son tipos demostrativos que derivan su referencia de una perspectiva de primera persona. Lo que examinará Loar es si estos aspectos de lo mental generan dificultades sustanciales a la tesis de que las cualidades fenoménicas son propiedades físicas.

Para Loar es fundamental subrayar la diferencia entre *conceptos* y *cualidades* (o propiedades), pues comparte la idea de Levine de que los argumentos anti-fisicalistas parten de sólidas intuiciones acerca de los *conceptos* pero van más allá, haciendo conclusiones acerca de las *cualidades*. La propuesta entonces es aceptar las intuiciones acerca de los conceptos (como, por ejemplo, que los conceptos fenoménicos son irreducibles), aceptando al mismo tiempo que las cualidades fenoménicas son idénticas a propiedades físicas. Loar señala que no pretende negar el poder de la intuición anti-fisicalista, sino dar una explicación de la relación entre conceptos fenoménicos y propiedades físicas que alivie esta “*ilusoria intuición metafísica*” (Loar, 1997, p. 598).

La argumentación empieza así con la observación común de que podemos saber que una persona tiene una propiedad física sin que sepamos que experimenta una cualidad fenoménica, y que ningún razonamiento *a priori* puede cerrar esta brecha conceptual. No

obstante, la conclusión que el anti-fisicalista extrae de esto (cf. Jackson, Chalmers) es que la cualidad fenoménica no puede ser idéntica con la propiedad física: si las concepciones física y fenoménica sólo pueden conectarse *a posteriori*, las propiedades físicas deben ser distintas de las fenoménicas.⁴⁰

Por supuesto, la premisa que esta línea de razonamiento presupone es que si dos tipos de propiedades son idénticas, debe poderse saber *a priori*, o contra-recíprocamente que si dos tipos de propiedades no están relacionados *a priori*, entonces no pueden ser idénticos. Esta presuposición es lo que Loar llama la “Premisa Semántica”.

La respuesta inicial de Loar es que el fisicalista podría aceptar que Mary no sabe que tenemos ciertas experiencias aunque sepa todos los hechos físicos sobre nosotros, e incluso que cuando adquiere dichas experiencias aprende algo nuevo sobre nosotros. Pero sólo aceptaría esta descripción en una lectura opaca y no en una lectura transparente del caso de Mary, pues para el fisicalista las propiedades fenoménicas son físicas (y entonces Mary ya conocía todas nuestras propiedades físicas, bajo sus descripciones físicas; lo que adquirió fue el conocimiento de tales propiedades en términos experienciales). Loar subraya que el principal defecto de argumentos como el de Mary es que extraer conclusiones metafísicas de contextos opacos es problemático, e introduce un par de contraejemplos que muestran cómo en los contextos opacos de lo que se trata muchas veces es de una nueva concepción de viejos hechos. Uno de los contraejemplos involucra un ‘modo de presentación’ descriptivo y otro científico, y el otro involucra un ‘modo de presentación’ científico y un ‘modo de presentación’ perceptual.⁴¹

⁴⁰ Estrictamente hablando, el anti-fisicalista podría decir que ningún razonamiento *a priori* ni *a posteriori* pueden cerrar esta brecha conceptual, pues el razonamiento *a posteriori* sólo puede vincular las concepciones física y fenoménica de manera contingente, sin dar cuenta satisfactoriamente de esta conexión.

Sin embargo, la intención de Loar parece ser el enfatizar la incapacidad *a priori* del fisicalista para vincular lo físico y lo fenoménico, más que la posibilidad de vincularlos *a posteriori*. Más adelante mencionaré que en la discusión que Ned Block y Robert Stalnaker (1998) hacen sobre el *Argumento del Análisis Conceptual* presentado por Jackson (1993) se ve la importancia que tiene el hacer este énfasis sobre la incapacidad *a priori* del fisicalista para vincular lo físico y lo fenoménico.

⁴¹ Los ejemplos son el de Max (quien sabe que la botella ante él contiene $\text{CH}_3\text{CH}_2\text{OH}$, pero no sabe que esa botella contiene ‘el ingrediente del vino y la cerveza que te intoxica’) y el de Margot (quien conoce el elemento oro y que la gente se decora con oro, pero nunca ha visto oro). Max carece de conocimiento de la sustancia bajo una descripción contingente del mismo y Margot carece de una cierta concepción visual del Au.

La línea de argumentación que el defensor del argumento del conocimiento podría elegir es la siguiente (basada en la *premisa semántica*). Los dos contraejemplos mencionados difieren del caso de Mary en un punto central: el conocimiento del que se carece en ellos está dado bajo un 'modo de presentación' contingente del fenómeno en cuestión (uno descriptivo y otro perceptual), por eso se puede decir que se trata de una nueva concepción de viejos hechos. Pero Mary no está concibiendo contingentemente una propiedad: ser experimentado de tal o cual manera es esencial para la propiedad que concibe.

Así, el defensor del argumento del conocimiento puede preguntar por qué el fiscalista diría este sólo sería válido en una lectura opaca, como una nueva concepción de viejos hechos, si no hay contingencia en la concepción que Mary tiene de la nueva información fenoménica (que la explique como una nueva concepción de viejos hechos).

En esta medida, la suposición de la argumentación anti-fiscalista es que la única forma de explicar el carácter *a posteriori* de una identidad de propiedades que vincula dos términos conceptualmente independientes,⁴² es que al menos uno de ellos seleccione la propiedad que refiere mediante un modo de presentación contingente. Pero, según Loar, esta suposición puede negarse.

Añadiendo esta consideración al argumento de la *premisa semántica*, dicho argumento puede refinarse de la siguiente manera:

Si dos tipos de propiedades son idénticas, debe poderse saber *a priori* (si dos tipos de propiedades no están relacionados *a priori*, entonces no pueden ser idénticos), *siempre que sus términos refieran de manera no-contingente*. Si una identidad de propiedades es *a posteriori* (pues sus términos son conceptualmente independientes), al menos uno de los términos debe seleccionar la propiedad que refiere mediante un modo de presentación contingente.

Para Loar, esta argumentación está emparentada con la de Kripke, quien asume (a) que un concepto fenoménico no puede estar ligado *a priori* con un concepto físico y (b) que una identidad puede ser verdad incluso si no lo es *a priori*. Lo que parece conjeturar el defensor del argumento de la *premisa semántica* es que Kripke asume (a) y (b) porque

⁴² Es decir, cuya verdad no se puede conocer *a priori* pues sus conceptos no se pueden vincular *a priori* (son lógicamente independientes, sólo se pueden vincular *a posteriori*).

piensa que lo que cuenta para el estatus *a posteriori* de la identidad es que haya involucrado un modo de presentación contingente.⁴³

Loar concluye señalando que la intuición anti-fisicalista que vincula la individuación de conceptos con la de propiedades procede de la siguiente manera: los conceptos fenoménicos y los físicos conciben sus referentes esencialmente, pero si dos conceptos conciben la misma propiedad *esencialmente*, uno debería poder ver *a priori* que seleccionan la misma propiedad. Lo que tiene que mostrar Loar, entonces, es que un concepto fenoménico puede seleccionar directa o esencialmente la misma propiedad que selecciona un concepto físico y ser *conceptualmente independiente* de todos los conceptos físicos (de manera que sustituir unos por otros en un contexto epistémico interpretado de manera opaca no preservaría la verdad). En este caso, nuestra concepción de las cualidades fenoménicas seguiría siendo directa/esencial, no-contingente, pero el fisicalismo no implicaría que conocer los hechos físicos implique conocer los hechos fenoménicos (el fallo en la implicación sería compatible con el fisicalismo).

2.4.2 LOS CONCEPTOS DE RECONOCIMIENTO

La solución de Loar empieza por plantear que los conceptos fenoménicos son una subclase de conceptos de reconocimiento. Los conceptos de reconocimiento son tipos demostrativos de la forma 'x es de esa clase', basados en disposiciones a clasificar fenómenos mediante ciertas discriminaciones perceptuales; no son descriptivos, no implican referencia a instancias pasadas, no dependen del acceso consciente a rasgos componentes, y son individuados en parte por la perspectiva desde la que se constituyen. La afirmación clave de Loar es que los conceptos fenoménicos son conceptos de reconocimiento que seleccionan propiedades físicas. Además, los

⁴³ Puesto lo anterior en términos del caso de Mary, el fisicalista estaría obligado a decir que una información sobre la experiencia de colores adquirida después de salir de la habitación es nueva si y sólo si involucra un modo de presentación contingente (i. e., una información que sea necesaria, que involucre un modo de presentación esencial, no puede ser novedosa, tiene que conocerse de antemano, *a priori*). Pero saber *cómo se ve el rojo* no involucra un modo de presentación contingente, luego no debería ser algo novedoso para Mary al salir de la habitación; el hecho de que lo sea indicaría que no puede tratarse de nada idéntico a algo que Mary conociera dentro de la habitación.

conceptos fenoménicos serían conceptualmente independientes de descripciones físicas y aún así seleccionarían (de manera no–contingente) estas propiedades; por esta independencia, los conceptos fenoménicos serían irreducibles a conceptos físicos, pues ni implican ni son implicados *a priori* por los conceptos físicos.

No obstante, rechazar la premisa semántica requiere justificar la idea de que los conceptos fenoménicos capturan propiedades físicas directa/esencialmente con independencia lógica de los conceptos físicos. ¿Cómo justificar esta afirmación de independencia conceptual? Recordemos que tanto en el caso de Max como en el de Mary tenemos conceptualmente independientes. Pero en el caso de Max, esta independencia conceptual se explica por un modo de presentación contingente asociado. Según Loar, en el caso de Mary la independencia conceptual se explicaría en la medida que los conceptos de ambas clases tiene roles conceptuales completamente diferentes: las habilidades básicas de reconocimiento asociadas con los conceptos fenoménicos no dependen o se disparan por el análisis científico consciente asociado con los conceptos físicos. Esto sugeriría un sentido en el que no es necesaria la independencia metafísica para la independencia conceptual de conceptos correferenciales (en este caso, de reconocimiento o teóricos).

De ser correcta la posición esbozada por Loar, implicaría que los juicios “el estado *a* se siente así” y “el estado *a* tiene la propiedad física/funcional P” pueden tener las mismas condiciones de verdad, incluso aunque su verdad o falsedad puedan conocerse sólo *a posteriori*. Un concepto fenoménico designaría rígidamente una propiedad, pero entonces designaría rígidamente la misma propiedad que designa un concepto físico.

De acuerdo con Loar, la objeción del antifisicalista inspirada en el argumento de la *premisa semántica*, podría ir en la siguiente dirección. Si un concepto designa rígidamente mediante un modo de presentación no–contingente, debe capturar la *esencia* de esa propiedad; pero si dos conceptos capturan esencialmente la misma propiedad ¿no deberíamos poder notarlo *a priori*? La respuesta de Loar es que tenemos aquí dos usos de ‘capturar la esencia’: uno equivale a ‘*designa rígidamente de manera directa*’, el otro significa ‘*es conceptualmente inter–derivable con algún predicado teórico*’.

que revela la estructura interna de la propiedad designada'. Pero, según él, la designación rígida no implica que los dos conceptos deban ser inter-derivables *a priori*.⁴⁴

Según Loar, de lo que se trata todo el asunto es simplemente que cuando ponemos juntos los conceptos fenoménicos y los conceptos físicos-teóricos, los estados cognitivos en los que nos ponemos al usarlos son tan diferentes fenomenológicamente que puede crearse la ilusión de que sus referentes deben ser diferentes:⁴⁵ “Es como si las intuiciones anti-fisicalistas descansaran en una teoría de la semejanza de la representación mental, como si concluyéramos a partir de la falta de semejanza en nuestras concepciones fenoménicas y físicas-funcionales una carencia de identidad en las propiedades a las que se refieren” (Loar, 1997, p. 605).

2.4.3 Las identidades *a posteriori* entre términos conceptualmente independientes

El argumento de la *Premisa Semántica* parece contener varios presupuestos sustanciales, que pueden cuestionarse al punto de socavar el argumento. Por un lado está el asunto de interpretar la expresión “designa rígidamente a *p*” como “captura la esencia de *p*”, en especial cuando más bien podría interpretarse “captura la esencia de *p*” como “designa rígidamente a *p*”: ante la dificultad de precisar qué es la esencia de un

⁴⁴ Si bien considero que esta respuesta de Loar cumple con su propósito de neutralizar la objeción del anti-fisicalista, es pertinente presentar algunos interrogantes a propósito de tal objeción. En ésta, asume la premisa semántica como tal: Si dos tipos de propiedades son idénticas, debe poderse saber *a priori* (siempre que sus términos refieran de manera no-contingente). Pero ¿cuál es el argumento filosófico que justifica la afirmación según la cual “si dos conceptos conciben la misma propiedad esencialmente, debe poderse saber *a priori* que seleccionan la misma propiedad” (es decir, que si lo hacen, debe poderse saber *a priori* la verdad de la identidad)?, ¿cómo justificar que se excluya la posibilidad de que pueda haber dos modos conceptualmente independientes de capturar la misma propiedad esencial de un objeto? Se trataría aquí del ‘sesgo cartesiano’ que Levine le atribuye a Kripke, pero llevado a sus más extremas consecuencias: si todo aquello que puede concebirse de manera clara y distinta puede el caso, entonces sólo puede ser el caso lo que puedo saber de manera clara y distinta (a través de razonamientos *a priori*). Como veremos más adelante, si este fuera el caso no podrían haber identidades entre modos de designación conceptualmente independientes (como de hecho sucede).

⁴⁵ De acuerdo con Loar, tal diferencia fenomenológica es explicada, al menos en parte, por Loar de la siguiente manera: el ejercicio de un concepto fenoménico en ausencia de la cualidad fenoménica asociada involucra ‘imágenes’, por lo que el concepto fenoménico tiene ‘afinidad fenomenológica’ con el estado fenoménico, afinidad fenomenológica que no se presenta cuando consideramos fenomenológicamente el concepto físico-teórico. Este argumento recuerda inevitablemente la distinción de Nagel entre la imaginación ‘perceptual’ y la imaginación ‘simpática’.

objeto, más bien podría entenderse la expresión “captura la esencia de p ” a través de la noción mucho más clara de “designa rígidamente a p ” (es decir, en todo mundo posible donde existe el referente, es designado mediante la expresión p).

Por otro lado, y de manera más importante, está la inferencia según la cual si dos conceptos capturan la esencia de una misma propiedad (es decir, la designan rígidamente), deberíamos poderlo saber *a priori*. En este caso, como correctamente advierte Loar se está interpretando la expresión ‘capturar la esencia’ al mismo tiempo como ‘*designa rígidamente*’ y como ‘*es conceptualmente inter-derivable con algún predicado teórico que revela la estructura interna de la propiedad designada*’. Sin embargo, la duda que surge de inmediato es que la designación rígida no parece contener la implicación de que los dos conceptos deban ser inter-derivables *a priori*.

En *El Nombrar y la Necesidad* (Kripke 1980, tercera lectura) Kripke es enfático en que el asunto de las propiedades esenciales de un objeto sólo se puede considerar correctamente si tenemos en cuenta la distinción entre aprioricidad (epistemológico) y necesidad (ontológico); es al segundo ámbito que corresponde la consideración sobre las propiedades esenciales del objeto. En este sentido, puede que no sepamos *a priori* que dos conceptos capturan esencialmente la misma propiedad, pero esto no necesariamente significa que no lo hagan; la designación rígida no implica entonces que los dos conceptos deban ser inter-derivables *a priori*.

Si, como parece admitirse, conceptos ‘Agua’ y ‘Calor’ son designadores rígidos, y los conceptos teóricos también lo son (se dice que ambos son términos de clase natural), se sigue que *ambos* capturan la esencia de su referente; pero esto es algo que claramente no sabemos *a priori* (pues descubrimos que el agua es H_2O y que el calor es la tasa de movimiento molecular). En este sentido, si aplicáramos lo señalado en el argumento de la *Premisa Semántica*, es decir, que una identidad *a posteriori* que vincula dos términos conceptualmente independientes es verdad sólo si al menos uno de ellos selecciona contingentemente su referente, no serían posibles las identidades *a posteriori* entre designadores rígidos: o bien todas tendrían que ser *a priori* (una identidad sólo sería informativa cuando hay contingencia), o bien todas serían falsas.

Ahora, si descartamos estos presuntos problemas en la idea de que puede existir identidades *a posteriori* entre términos conceptualmente independientes, vale decir lo siguiente: que la noción de *conceptos de reconocimiento* adelantada por Loar parece compatible con la idea de que el modo de concebir subjetivo de primera persona es conceptualmente independiente del modo de concebir objetivo de tercera persona de la ciencia (idea fundamental para afirmar por qué no puede cerrarse la brecha explicativa). Es imposible encontrar el tipo de conocimiento que se requiere para cerrar la brecha explicativa, pues no podemos conciliar los siguientes requerimientos en un único conocimiento: extensionalidad, descriptividad, perspectiva de tercera persona y, a la vez, subjetividad, demostratividad que derive su referencia de una perspectiva de primera persona. Quizás la única sombra de duda que podría arrojarse es si los conceptos de reconocimiento se pueden adquirir sin haber tenido la experiencia. Pero esto parece incompatible con su carácter demostrativo, derivado de su referencia a una perspectiva de primera persona: en este sentido, sólo quien ha tenido la experiencia puede comprender el concepto fenoménico en cuestión, lo cual excluye la posibilidad de que se pueda comprender un concepto fenoménico sin haber tenido la experiencia asociada.

2.4.4 LA BRECHA EXPLICATIVA SEGÚN LOAR

Un paso obligado en la discusión que lleva a cabo Loar es la valoración de la brecha explicativa. En este respecto, admite una concepción de explicación cercana a la del modelo de reducción de Levine, ya que concede que cuando explicamos la liquidez en términos físicos–funcionales la explicación es *en parte a priori*, pues analizamos *a priori* los aspectos de la liquidez en términos de una descripción funcional y *luego* mostramos que la teoría física del agua implica *a priori* que la descripción funcional es realizada.⁴⁶

⁴⁶ El razonamiento que Loar está concediendo aquí es una formulación de lo que Block y Stalnaker (1998) llaman el argumento del *Análisis Conceptual*, defendido por Jackson (1993). Como he mencionado, sobre la base de escenarios tipo tierra–gemela, entre otros, Block y Stalnaker (1998) rechazan la posibilidad de hacer esta derivación *a priori* que incluye premisas empíricas, incluso en el caso del agua. Por supuesto, para Jackson y Chalmers, el argumento del análisis conceptual tiene claras consecuencias ontológicas y no meramente epistemológicas como supone Loar.

Por otro lado, la diferencia con la concepción de Levine tiene que ver con la etapa 2 de su modelo de reducción, en la que se realiza el trabajo *empírico* de descubrir los mecanismos que realizan la descripción funcional de la etapa 1. Dentro de lo que Loar concede, la segunda etapa también adopta una teoría empírica (la teoría física del agua), pero lo clave en ella es que se muestra que dicha teoría empírica implica *a priori* que la descripción funcional es realizada. Parece posible

Pero dada la independencia conceptual de los conceptos fenoménicos y los físicos/funcionales, no podemos tener tal explicación *a priori* de las cualidades fenoménicas en términos físicos/funcionales⁴⁷. No obstante, para él esto únicamente revela que la brecha explicativa es un fenómeno epistémico o conceptual, que no tiene consecuencias ontológicas y que puede predecirse a partir de su explicación fisicalista.

Loar reconoce que esta forma de despachar la brecha explicativa puede dejarnos una sensación extraña. De acuerdo con él, no puede ser la simple independencia lógica entre ciertos conceptos de reconocimiento demostrativos y ciertos conceptos descriptivos de la ciencia física la que genera una brecha explicativa (como lo muestra el que no encontremos brecha en juicios como “esa sustancia es $\text{CH}_3\text{CH}_2\text{OH}$ ”, incluso aunque no sea *a priori*). ¿Qué genera entonces la brecha explicativa?

La respuesta que ofrece Loar es que la brecha se genera porque suponemos que los enunciados de identidad verdaderos *deberían* ser explicativos, pero los enunciados de identidad que conectan conceptos fenoménicos y conceptos físicos/funcionales no lo son. Pero hemos encontrado enunciados de identidad que pueden no ser explicativos (por su independencia conceptual) y ser verdaderos, así que lo que tenemos que explicar más bien es la intuición de que todo enunciado de identidad debería ser explicativo.

Como lo ha evidenciado el argumento de la ‘premisa semántica’, la intuición se debe a que (a) consideramos que los conceptos fenoménicos capturan sus cualidades como son en sí mismas (es decir esencial o directamente, no contingentemente), y (b) también pensamos que los conceptos físicos/funcionales capturan esencialmente la propiedad que seleccionan; así que llegamos a creer que (c) si la identidad fuera verdadera, debería sernos evidente la mismidad de esa propiedad a partir de esas concepciones. El que esto no suceda nos dificulta entender cómo puede haber una sola propiedad ahí. Pero de hecho es predecible que no suceda, por el tipo de independencia conceptual que se da

llevar la concepción de Levine a una convergencia con la posición que describe Loar, pues el énfasis que el segundo hace en la etapa 2 es compatible, y podría añadirse, a la manera en que el primero concibe esta etapa.

Véase más adelante el apartado 7.3.1.

⁴⁷ Decir que no podemos tener una explicación física o funcional de las cualidades fenoménicas *porque* los conceptos fenoménicos son independientes de los conceptos físicos o funcionales, puede producir algo de recelo. Que no podamos tener semejante explicación, en efecto, se deriva de dicha independencia conceptual, pero esta, a su vez, es resultado de la naturaleza especial de las experiencias conscientes (que es lo que ha llevado a muchos a pensar que no son entidades físicas). Como veremos, esta claridad resultará relevante unas líneas más adelante cuando Loar comenta el enfoque de Nagel.

entre ambos términos: la idea de que los conceptos fenoménicos son conceptos de reconocimiento de cierta clase (que no conciben sus referentes contingentemente) explica la brecha de manera compatible con el fisicalismo.⁴⁸ Como señala Loar:

El problema de la brecha explicativa proviene entonces de una ilusión. Lo que genera el problema es el no apreciar que puede haber dos ‘aprehensiones directas’ independientes conceptualmente de una misma esencia, es decir, aprehensión demostrativa al experimentarla, y aprehensión en términos teóricos. La ilusión es de una transparencia esperada: una aprehensión directa de una propiedad debería revelar cómo está constituida internamente, y si no se revela como físicamente constituida entonces no lo es. El error es la idea de que una aprehensión directa de esencia debería ser una aprehensión transparente (...). (Loar, 1997, p. 609).

2.4.5 La Brecha Explicativa y Argumento del Análisis Conceptual: Levine, Loar, Chalmers, Jackson, Block y Stalnaker.

En la medida que Loar no sólo describe sino que también concede la posición del ‘análisis conceptual’, se sigue por lo que he dicho que existe alguna relación entre la noción de explicación–reducción de Levine–Loar y la de Chalmers–Jackson. En efecto, ambos comparten la siguiente premisa: que los conceptos fenoménicos ni implican ni son implicados *a priori* por los conceptos físicos, que son irreducibles, que son conceptual o lógicamente independientes; en suma, comparten la idea de que hay una brecha explicativa entre la experiencia consciente y la actividad cerebral. La diferencia reside en

⁴⁸ Una réplica posible se basa en el caso de identidades entre términos conceptualmente independientes, donde pese a la independencia conceptual la identidad es explicativa, como la identidad “agua=H₂O”. La respuesta de Loar es que esta es una forma completamente diferente de captar la esencia de la que tenemos en el caso de los conceptos fenoménicos, pues el término “agua” puede hacerse conceptualmente equivalente a alguna descripción funcional que se deriva del término teórico de la identidad. Esta respuesta puede resultar problemática en la medida que no es completamente claro que términos como “agua” sean equivalentes a descripciones (y esto es precisamente lo que fundamenta la idea de que el término “agua” es conceptualmente independiente del término “H₂O”). Sin embargo, puede respaldarse la idea de que términos como “agua” capturan la esencia de sus referentes de una forma completamente diferente a como lo hacen los conceptos fenoménicos, pues captura referentes dados objetivamente (en el sentido de ser públicos, observables) mientras que los conceptos fenoménicos captura referentes dados subjetivamente. Es un error pensar que, si el fisicalismo es verdad, las identidades que involucran la experiencia consciente deben funcionar como las identidades que involucran términos como “agua”.

que para la primera pareja esto no tiene consecuencias ontológicas, mientras que para la segunda dupla sí las tiene.

Sin embargo, Block y Stalnaker (1998) atacan vigorosamente tanto el modelo de reducción de Levine como el argumento del análisis conceptual. De acuerdo con estos dos filósofos, es central para una defensa del fisicalismo la idea de que no hay un buen argumento que demuestre la asimetría entre las identidades agua=H₂O y estados cerebrales=experiencias conscientes. En la medida que la afirmación de asimetría no ha sido apoyada apropiadamente, no hay ninguna razón por la que no podamos aceptar el fisicalismo (además de que, si la hubiera, posiblemente la segunda identidad mencionada no pasaría el 'test' propuesto por Jackson –y compartido por Chalmers– para justificar la afirmación de necesidad). En resumen, para Block y Stalnaker no se ha mostrado que haya una brecha (incluso es posible que no la haya).

La posición que he venido defendiendo seguramente esté más próxima a la del fisicalismo Levine–Loar que a la del fisicalismo Block–Stalnaker. ¿Cómo responder entonces a la acusación de que no se ha mostrado que haya brecha explicativa? Pienso que buena parte de esta pregunta puede encontrarse en mi discusión sobre Levine, en la noción de derivabilidad que examino allí, aunque es probable que se replique (en una vena como la de Block y Stalnaker) que esta noción de derivabilidad no solamente no puede aplicarse para la identidad estados cerebrales=experiencias conscientes sino que tampoco puede aplicarse realmente a la identidad agua=H₂O. Habría que ver si a partir de esta noción de derivabilidad se puede encontrar aún un sentido en el cual las identidades no sean simétricas.

Ahora bien ¿cómo responder a la acusación de que, si hubiera asimetría, la identidad estados cerebrales=experiencias conscientes no pasaría el 'test' propuesto por Chalmers y Jackson para justificar la afirmación de necesidad? Creo que esta acusación es más fácil de capotear para el fisicalista que admita la existencia de la brecha, pues podría cuestionarse la idea de que ese sea un test válido o justificado de necesidad; podría afirmarse que es una idea que pone un asunto de ontología (la necesidad) en términos de un asunto epistemológico (la derivabilidad a priori) a través de la noción de justificación (sin mencionar que podrían exponerse también las dudas de Block y Stalnaker acerca de si esta derivación de hecho puede hacerse en otros casos; aunque

para arrojar estas dudas sin sacrificar la idea de que hay una brecha explicativa, es necesario haber articulado una idea de asimetría compatible con tales dudas).

Quepa decir, para terminar, que la posición de Loar por momentos no parece comprometida tan claramente con la noción de brecha explicativa de Levine:

Como apunta Georges Rey (...), el mero hecho de la inequivalencia conceptual para los tipos–demostrativos de reconocimiento y los términos de descriptivos no genera una brecha explicativa. (...) No encontramos una problemática brecha explicativa en juicios de la forma “esa sustancia es $\text{CH}_3\text{CH}_2\text{OH}$ ”, incluso aunque esto no se sostenga *a priori*. (Loar, 1997, p. 608)

Loar parecería pensar aquí que la independencia lógica, conceptual, no es suficiente para crear una brecha explicativa. Si fuera suficiente, deberíamos pensar que hay una brecha explicativa en enunciados como “esa sustancia es $\text{CH}_3\text{CH}_2\text{OH}$ ” o “esa sustancia es H_2O ” (o incluso, si se acepta que expresiones como ‘agua’ y ‘alcohol’ tiene cierto componente demostrativo, enunciados como “agua es H_2O ” y “alcohol es $\text{CH}_3\text{CH}_2\text{OH}$ ”); en los enunciados anteriores se vinculan tipos–demostrativos y términos descriptivos, sin que haya brecha alguna. ¿Por qué entonces creemos que sí hay una brecha cuando relacionamos otros tipos–demostrativos (los conceptos fenoménicos) y otros términos descriptivos (los conceptos físicos)?

En otras palabras: si fuera sólo por la independencia lógica, no habría brecha explicativa, dicha independencia es insuficiente para afirmar una brecha (precisamente lo que aseveran Block y Stalnaker). Loar deduce que debe ser otra la razón por la cual pensamos que hay una brecha entre los conceptos fenoménicos y los conceptos físicos; propondrá que la se trata del hecho de que la relación estados cerebrales=experiencias conscientes *no es explicativa, como esperamos que fuera*. ¿Significa esto que hay una brecha explicativa o que *pensamos* que la hay? La ambigüedad en la argumentación de Loar dificulta verdaderamente responder esta pregunta; plantear que no hemos reflexionado alrededor de que la misma esencia puede capturarse de dos modos conceptualmente independientes no resuelve el problema: podría significar tanto que esa omisión hace ilusoria la brecha como que la hace real aunque inofensiva.

2.4.6 Nagel y el juego de “dejar por fuera”

Loar suscribe el planteamiento de Nagel según el cual lo mental sólo es accesible desde un punto de vista, pero Loar piensa que esto no refleja algo acerca de los estados fenoménicos sino acerca de los conceptos fenoménicos: *“Nagel hace una observación correcta sobre los conceptos y extrae una conclusión errónea sobre los hechos y propiedades. Pues los conceptos pueden en ese sentido ser “desde un punto de vista” y subjetivos, y aún así introducir propiedades que son capturadas exhaustivamente en la ciencia objetiva”* (Loar, 1997, p. 609)

Realmente no sé de dónde provenga la subjetividad de los estados fenoménicos, no sé por qué son “desde un punto de vista” (de hecho creo que nadie lo sabe), pero es absurdo pensar que los estados fenoménicos no son subjetivos. La subjetividad de los conceptos fenoménicos es derivativa de la subjetividad de los estados fenoménicos a los que dichos conceptos se refieren (los conceptos fenoménicos son “desde un punto de vista” porque los estados también lo son). El problema que parece estar enfrentando Loar, y que resuelve diciendo que son los conceptos y no los estados fenoménicos los que son subjetivos, es uno bastante gastado: si los estados fenoménicos son esencialmente subjetivos ¿cómo pueden los referentes de los conceptos fenoménicos ser entidades físicas objetivas? Resuelve proponer que esta subjetividad resulta derivada de los conceptos pre-científicos que utilizamos para capturar estos referentes, pero que estos referentes pueden ser capturados cabalmente de una manera no subjetiva: *“los conceptos pueden en ese sentido se “desde un punto de vista” y subjetivos, y aún así introducir propiedades que son capturadas exhaustivamente en la ciencia objetiva”* (Loar, 1997, p. 609). Esta idea se repite incluso cuando admite que son subjetivos los referentes y no los conceptos: *“¿Por qué no debería esta subjetividad misma ser idéntica con una propiedad física-funcional, y por ende completamente concebible objetivamente bajo su descripción física-funcional? No hay contradicción en suponer que una propiedad que es subjetiva (...) también es concebible bajo un modo de presentación objetivo”* (Ibíd.).

Loar parece haber caído presa de lo que él mismo denomina *“un juego con ‘dejar por fuera’* (Loar, 1997, p. 610), en el que sus afirmaciones muestran una confusión entre lo que Levine llama la “lectura epistemológica” y “la lectura ontológica” de ‘dejar por fuera’: en la primera lectura, afirmaciones como “las propiedades fenoménicas son capturadas

exhaustivamente en la ciencia objetiva”, o “la subjetividad es completamente concebible objetivamente bajo su descripción física-funcional” son falsas, aunque puedan ser verdaderas bajo la segunda lectura. Lo que he planteado es que son falsas en la lectura epistemológica y pueden ser verdaderas en la lectura ontológica por la siguiente razón: desde cierta perspectiva epistémica o modo de concebir (el del sujeto que los tiene, en ‘primera persona’) se trata de estados subjetivos, pero desde otra perspectiva epistémica o modo de concebir (el del observador que los aprecia, en ‘tercera persona’) se trata de estados neuronales. Por esta razón, en cuanto experiencias conscientes (un modo de concebir), la ciencia (otro modo de concebir) no va a poder capturar nunca estas propiedades. Como señala el mismo Loar:

¿Una descripción completamente objetiva de la realidad no deja aún algo por fuera, a saber, las concepciones subjetivas? Este es un juego con ‘dejar por fuera’ Una descripción objetiva completa deja por fuera concepciones subjetivas, no porque no pueda caracterizar completamente las propiedades (...) sino simplemente porque no las emplea. (Loar, 1997, p. 610).

2.4.7 Acerca de Internalismo vs Externalismo

El hecho de que diferenciamos nuestras experiencias conscientes por *cómo se siente tenerlas*, se relaciona directamente con el largo pasaje que al comienzo del artículo desarrolla Loar acerca del debate entre una interpretación externalista y una internalista de ‘cualidad fenoménica’. En la interpretación externalista, ‘cualidad fenoménica’ hace referencia a la manera que tiene de aparecer *el objeto* de la experiencia, por lo que se adscribe ‘cualidad fenoménica’ al objeto y no a la experiencia. En la interpretación internalista que reconstruye Loar (la cual resulta muy afín la fenomenología Husserliana), las experiencias conscientes no involucran intrínsecamente relaciones con cosas fuera de la mente, sino que tienen una estructura *intencional* intrínseca (por la cual el hecho de que la experiencia nos presente las cosas como siendo de tal o cual manera es un rasgo no-relacional de la misma).

Así, en la interpretación internalista de ‘cualidad fenoménica’, decimos que concebimos cómo difieren entre sí las experiencias por algo no-relacional, por una propiedad intrínseca de las experiencias: por *cómo se siente tenerlas* (en inglés, *what it is like to have them*). Esto es, lo que concebimos mediante los conceptos fenoménicos son las

cualidades de las experiencias (independientemente de cualesquiera atribuciones que hagamos a las cualidades de los objetos de la experiencia).

Loar señala que privilegia la interpretación internalista, en la medida que piensa que es la que adhiere el anti-fisicalista –su blanco en este artículo–. Una consideración exhaustiva de estas posiciones y del debate entre internalistas y externalistas rebasa con mucho los propósitos de este escrito, sin embargo, vale la pena hacer un brevísimo excursus sobre la naturaleza y dimensiones del debate internalismo/externalismo acerca de la experiencia consciente.

Dicho de una manera esquemática, el externalismo es la posición según la cual, para tener ciertos tipos de estados mentales intencionales (creencias, deseos, etc.), es necesario relacionarse con el ambiente de la manera correcta; el internalismo, por su parte, sostiene que tener estados mentales intencionales depende solamente de propiedades intrínsecas de los estados mentales. Ahora, más allá de mantener una posición externalista con respecto a los estados mentales conocidos usualmente como ‘epistémicos’ o ‘cognitivos’ (creencias, deseos), algunos han llegado a sostener esta postura con respecto a las experiencias conscientes.⁴⁹ En efecto, en virtud de lo que se ha llamado el *Argumento de la Transparencia* (Harman 1990), algunos teóricos (Harman, Burge, Tye, entre otros) han llegado a la conclusión de que reflexionar o hacer introspección sobre cómo es tener una experiencia no muestra que seamos conscientes de propiedades intrínsecas de las experiencias mismas, sino solo de las propiedades de los objetos experimentados. La versión más radical del argumento de la transparencia sostiene que ni siquiera la introspección más cuidadosa relevaría propiedades intrínsecas de las experiencias mismas, pues estas propiedades simplemente no existen.

El externalismo acerca de la experiencia consciente ha sido controvertido, pues se ha replicado que es problemático suponer que el carácter fenoménico, cualitativo o subjetivo de la experiencia se agota en la naturaleza de los objetos y cualidades reales experimentadas. Lo típico es que las objeciones al argumento de la transparencia tengan la forma de contraejemplos, que pretenden mostrar que se puede hacer introspección de ciertos rasgos de nuestras experiencias aunque no sean representacionales. Block

⁴⁹ Por supuesto, esto supone que se haya dado otra discusión: si las experiencias conscientes son o no intencionales, lo cual supondremos *por mor* de la discusión.

menciona casos de sensaciones corporales o estados de ánimo cuyo contenido representacional parece ser nulo, pero de los que sin duda se puede hacer introspección. La respuesta del externalista es que incluso en estos casos hay un residuo, aunque sea mínimo de contenido representacional. Por esto, resulta más interesante otro argumento de Block, según el cual una escena visual puede parecer muy diferente si uno se quita los anteojos (mi experiencia visual se vuelve borrosa, los contornos de los objetos se vuelven distintos, etc.); esta diferencia fenoménica no parece derivarse de ninguna diferencia en los objetos de experiencia, sino que parece ser una diferencia en la manera en la que esos objetos son experimentados. Block nota que si mi visión se vuelve borrosa, se puede hacer introspección del carácter borroso, sin mencionar que se puede hacer introspección del carácter visual de la representación (en lugar de olfatorio), entre otras características que no se explican mediante diferencias en los objetos experimentados; por lo tanto, la introspección parece revelar un discernimiento de propiedades de las experiencias mismas tanto como de sus objetos (para ver una interpretación externalista de este fenómeno, véase Tye 2000).

Una consecuencia absolutamente relevante de este externalismo acerca de la experiencia consciente, es que estos teóricos se muestran reticentes a la afirmación de que los estados y eventos mentales son eventos y estados neurofisiológicos. No obstante, la clave es que no lo hacen por un rechazo de la metafísica fisicalista, sino porque adhieren una explicación externalista de las condiciones de individuación de los estados mentales. Para el externalista, la identidad de un estado mental está ligada a la identidad y existencia de objetos y propiedades en el mundo acerca de los cuales es el estado. Como señalé hace un momento, la consideración del debate entre internalistas y externalistas rebasa con mucho los propósitos de este escrito, así baste por ahora con la siguiente sugerencia que requerirá ser desarrollada en otra ocasión: el argumento básico del fisicalismo podría seguir corriendo, sólo que con la cláusula (que no sería necesariamente puro escrúpulo) de que además de las propiedades intrínsecas de los estados mentales, para la especificación de su contenido deberán tenerse en cuenta las propiedades relacionales de los mismos. Las propiedades intrínsecas de los estados mentales serían aquellas que se capturan, por un lado, en el modo de concebir subjetivo de primera persona, y, por otro lado, en el modo de concebir de tercera persona de la ciencia; por su parte, las propiedades relacionales serían, evidentemente, las relaciones con los ítems del mundo acerca de los cuales es el estado mental. De esta manera, en

un sentido, un estado mental sería una compleja entidad relacional, que incluiría tanto sus propiedades relacionales como sus propiedades intrínsecas (las cuales pueden ser capturadas desde dos modos distintos de concebir).

2.5 PAPINEAU: TENIENDO CUIDADO CON LA BRECHA

En apartados anteriores he anticipado algunos elementos de la posición de Papineau (Papineau, 1998). He subrayado que para él interrogantes como el que presenta Chalmers en su *problema difícil* (¿cómo es que los sistemas físicos son sujetos de experiencia?) o como otro al que alude Nagel (¿cómo es que a y b son idénticos?) constituyen preguntas fuera de lugar; lo anterior resulta del hecho de que está completamente fuera de lugar exigir que una identidad sea explicativa. La necesidad de explicar una identidad, de acuerdo con Papineau, sería improcedente. Como observé a propósito de las varias formulaciones del *problema difícil*, Papineau alega que a este tipo de preguntas no podemos responder sino con la afirmación de que así es el mundo: si $a=b$, entonces simplemente es así, no cabe preguntar por qué son idénticos.

De lo anterior se infiere que el supuesto desafío incontestable que establecería para el fisicalismo la existencia de una brecha explicativa, o la imposibilidad de responder a lo que llamé *la pregunta C*, sería espurio. No se podría afirmar que el fisicalismo tiene dificultades para explicar por qué los estados conscientes son idénticos a estados cerebrales, pues el fisicalismo es una tesis de identidad, y las necesidades no necesitan explicarse. Explicar cómo es que las experiencias conscientes son idénticas a ciertos estados cerebrales, dice Papineau, es buscar algo que no se necesita para admitir el fisicalismo.

2.5.1 El principal argumento a favor del fisicalismo: el argumento causal

Dado que las identidades no necesitan ser explicadas, cuando mucho lo que se requiere es mencionar las razones por las cuales pensamos que la identidad es el caso; esto es, necesitamos mencionar los argumentos por los cuales pensamos que el fisicalismo es verdad. Ante la pregunta de qué es lo que muestra que son idénticos, Papineau reconstruye tres posibles líneas de argumentación que podrían ser utilizadas como defensa del fisicalismo.

La primera, que sostiene que el fisicalismo está respaldado por una intuición directa, es descartada por Papineau al considerar que nuestras intuiciones son insuficientes para establecer una teoría (lo cual puede verse en el hecho de que tenemos intuiciones contradictorias); la segunda afirma que nuestros conceptos de estados conscientes están asociados *a priori* con ciertas descripciones funcionales, pero el análisis conceptual *a priori* no revela realmente estas asociaciones. La tercera línea de argumentación señala que, como cuestión de hecho, los estados conscientes y los estados físicos siempre aparecen en el mismo lugar en el esquema causal de las cosas y, por ende, deben ser idénticos. Esta línea de razonamiento toma dos premisas comúnmente admitidas; que las experiencias conscientes causan conductas y el de la *completitud de la física*, para construir el siguiente argumento:

1. Los eventos mentales conscientes tienen efectos físicos
2. Todos los efectos físicos tienen causas físicas suficientes

De las premisas anteriores se seguiría que la conducta tendría dos causas suficientes distintas. Pero independientemente de si hay o no eventos sobredeterminados, Papineau señala que este no parece el modelo correcto de causación mental: la sobredeterminación implica que incluso si una de las causas estuviera ausente, el resultado aún habría ocurrido por la otra causa (lo cual parece erróneo en el caso de la conducta). Así que se puede añadir la siguiente premisa.

3. Los efectos físicos de causas conscientes no siempre están sobredeterminados

¿Qué sucede entonces con las causas conscientes de los efectos físicos que no están sobredeterminados si se tiene en cuenta la completitud de la física? Deben ser idénticos a causas físicas. Esto permite admitir tanto 1 como 2 y evitar la implicación de sobredeterminación, pues no se trataría de causas distintas.⁵⁰

⁵⁰ Si se pregunta por la manera en que se están concibiendo los elementos de la relación causal, señala Papineau, da lo mismo si se los concibe como hechos, estados o Kim-eventos (instanciaciones particulares de propiedades), pues la identificación de las causas conscientes con las causas físicas requiere que la propiedad consciente sea idéntica con la propiedad física (los estados o hechos relevantes no serían idénticos si las propiedades no lo son). Sin embargo, anota Papineau, si se los concibe como Davidson-eventos, la identidad de Davidson-eventos conscientes con Davidson-eventos físicos sólo requeriría que las propiedades conscientes y físicas relevantes se instancien en el mismo particular (no que las propiedades sean idénticas); no se seguiría entonces la identidad de propiedades conscientes y físicas. Con todo, el argumento se puede rephrasing:

La forma más natural de atacar el argumento causal es tratar de probar que alguna de sus premisas no es obligatoria. No obstante, Papineau resalta que las posiciones que se derivan de negar alguna de estas premisas son mucho menos plausibles que la identidad de las propiedades conscientes y físicas: el rechazo de la primera premisa conduce al epifenomenalismo o al paralelismo psicofísico; el rechazo de la segunda premisa, conduce a negar que lo físico es causalmente completo (añade que si bien no es incoherente pensar que la experiencia es físicamente fundamental, va en contra de una rica evidencia empírica). Por su parte, el rechazo de la tercera premisa implicaría renunciar a nuestra concepción común de la causalidad, aceptando masivamente la sobredeterminación causal.

En resumen, ante la pregunta de qué es lo que muestra que los estados cerebrales y las experiencias conscientes son idénticos, se pueden responder que siempre aparecen en el mismo lugar en el esquema causal de las cosas.

2.5.2 ¿Puede el fisicalismo salirle al paso a los reparos del funcionalista?

Papineau no desconoce la intuición según la cual lo mental parece múltiplemente realizable, intuición que robustece las aproximaciones funcionalistas y que constituiría un sentido en el cual la identidad entre experiencias conscientes y estados cerebrales no podría sostenerse. Señala que el movimiento estándar del fisicalista es identificar los estados conscientes ya no directamente con estados cerebrales sino con cualquier propiedad de segundo orden (funcional) que sea común a los sustratos físicos de los organismos que tengan dichos estados conscientes. La pregunta que surge, por supuesto, es ¿estamos respetando las premisas del argumento causal?

1'. Todos los eventos conscientes causan algunos eventos físicos *en virtud de* sus propiedades conscientes.

2'. Todos los eventos físicos son causados por eventos físicos previos *en virtud de* las propiedades físicas de éstos.

3'. Los efectos físicos de las causas conscientes no siempre están sobredeterminados.

Para hacerlo consistente, identificamos las propiedades conscientes de las causas con sus propiedades físicas.

Estrictamente hablando, las propiedades conscientes no serían idénticas a propiedades físicas, sino a propiedades funcionales, que son a su vez realizadas por esas causas físicas. Los estados conscientes no causarían directamente los efectos conductuales, lo harían indirectamente a través de su sustrato físico. Pero si bien las experiencias no serían estrictamente físicas sino funcionales, *requerirían* sustrato físico, *requerirían* estar instanciados en algún 'hardware' físico para tener poder causal indirecto (de hecho, no podrían existir sin ese sustrato físico).⁵¹ Lo que se estaría haciendo, entonces, sería identificar los estados conscientes con estados de segundo orden *que necesitan estar físicamente realizados*.

Papineau plantea entonces que incluso en este caso el argumento causal se podría reformular. No obstante, esto requiere que aceptemos una noción más amplia de causalidad, en la que un estado puede causar (*indirectamente*) en virtud de tener un realizador que causa (*directamente*). Así, el defensor del argumento causal podría decir que los estados mentales causan efectos físicos en el sentido de que o bien los causan directamente o bien lo causan indirectamente (a través de realizadores que los causan directamente).

Podría objetarse que realmente tenemos dos causas aquí (el estado funcional y su realizador),⁵² pero ni siquiera admitiendo esto podría decirse que estas dos causas sobredeterminan el resultado, pues la propiedad funcional se presenta sólo en virtud de la presencia de la propiedad física. El resultado no podría haberse presentado si la propiedad física estuviera ausente, pues la propiedad funcional también lo habría estado.⁵³

⁵¹ Muy a la manera en que Aristóteles, al hablar de *forma* y *materia*, señala que las dos no pueden separarse sino 'en el análisis', podemos decir que aunque las propiedades o estados funcionales pueden pensarse separadamente de su realización física, no pueden existir sin *ningún* sustrato físico. Tales estados o propiedades de segundo orden no son entidades, existencias, independientes; no tienen una existencia propia independiente de la del sustrato físico en el que se dan (y, de manera importante, al cual no se les puede reducir explicativamente).

⁵² Véase el pie de página anterior.

⁵³ Esta versión reformulada del argumento causal es diferente de lo que Papineau ha llamado el 'argumento de la realización': dicho argumento señala que hay una conexión *a priori* entre conceptos de estados conscientes y descripciones funcionales, pero la reformulación del argumento causal asume que la identidad entre propiedades conscientes y funcionales físicamente instanciadas se deriva de premisas *a posteriori* (no se requiere ninguna conexión a

De esta manera, lo que nos propone Papineau es una especie de fisicalismo debilitado y compatibilizado con el funcionalismo⁵⁴, que si bien implica que el primero renuncie a algunas ciertas pretensiones (como la identidad *estricta*, e incluso a una noción cabalmente eficientista de la causalidad) básicamente podría mantener lo más básico de su posición inicial (sigue siendo verdad que todo lo que existe es o bien físico, o depende de lo físico para su existencia).

2.5.3 La Brecha según Papineau

¿Qué es lo que hace, entonces, que tantos filósofos sientan que hay una brecha explicativa?, pregunta Papineau. Su respuesta es que estos teóricos piensan que hay un sentido en el que algunas identidades pueden explicarse: cuando un enunciado de identidad tiene una descripción en uno de sus lados, podemos pedir una explicación de por qué la entidad satisface la descripción. Algo semejante sucede en el caso de los enunciados de identidad que incluyen términos científicos: debe de explicarse por qué la entidad teórica en cuestión satisface las descripciones asociadas *a priori* con el concepto cotidiano relevante. Tomando este ejemplo como paradigma, sienten que –si los estados conscientes y los físicos son idénticos– el fisicalismo debería estar en capacidad de explicar que ciertos estados físicos satisfagan ciertas descripciones que están asociadas *a priori* con nuestros conceptos de estados conscientes, pero –señalan– *por principio* estas explicaciones no se pueden obtener. La ausencia de estas explicaciones se ve como un obstáculo que bloquea cualquier acceso epistemológico al fisicalismo, como una brecha explicativa.

priori). La reformulación apunta a señalar que si las propiedades conscientes no son estrictamente idénticas con propiedades físicas, deben ser idénticas a propiedades funcionales físicamente realizadas (de otro modo no podríamos ver cómo causan sus efectos físicos). La reformulación no asume nada sobre los conceptos de propiedades conscientes ni sobre su conexión con descripciones funcionales; independientemente de la naturaleza de nuestros conceptos de propiedades conscientes, se referirían a propiedades funcionales físicamente realizadas.

⁵⁴ Creo, sin embargo, que la distancia entre el funcionalismo y el fisicalismo puede resultar menor de lo que usualmente se piensa. David Lewis (1971, 1980), por ejemplo, ha interpretado la identificación funcional de los estados mentales como algo que proporciona la base para una identificación subsiguiente de qué estados físicos los instancian. Cuando encontramos un estado físico que instancia el rol caracterizado funcionalmente, entonces tenemos establecido que es el mismo que el estado identificado funcionalmente.

Sin embargo, alega Papineau, esto no muestra nada en contra del fisicalismo (para empezar, lo que estamos haciendo aquí no es *explicando una identidad*, sino explicando por qué se satisface cierta descripción); sólo se muestra que nuestros *conceptos* de estados conscientes no están asociados *a priori* con ninguna *descripción* tal.

Papineau responde que carecemos de las mencionadas explicaciones, pero que este no es un problema para el fisicalismo porque no existen descripciones asociadas *a priori* con nuestros conceptos cotidianos de estados conscientes. En efecto, podemos mostrar que el “H₂O” satisface ciertas descripciones asociadas *a priori* con “agua” y de allí pasar a afirmar que el agua es H₂O; pero este patrón de argumentación es adecuado cuando nuestros conceptos pre-teóricos están asociados *a priori* con descripciones funcionales, pero no hay necesidad de insistir en esta forma de identificación cuando nuestros conceptos no están asociados *a priori* con descripciones funcionales (y no hay descripciones funcionales asociadas *a priori* con nuestros conceptos cotidianos de estados conscientes, por lo cual no se abriría la brecha).⁵⁵

Con todo, si nuestros conceptos de estados conscientes, nuestros conceptos fenoménicos, no están asociados *a priori* con descripciones ¿cómo refieren entonces? Antes de responder este interrogante es importante tener en cuenta, señala Papineau, a *qué* se refieren nuestros conceptos fenoménicos. Incluso si los conceptos fenoménicos no tuvieran conexión *a priori* con descripciones, no se sigue que no se refieran a propiedades materiales.

Papineau piensa que no es necesario admitir la postura según la cual los conceptos fenoménicos se refieren a propiedades fenoménicas distintas de cualesquiera propiedades físicas. En dicha postura, los conceptos fenoménicos se derivan de la aprehensión de primera mano de alguna propiedad fenoménica.

La propuesta de Papineau tiene como punto de partida la siguiente idea: el despliegue de conceptos fenoménicos involucra una especie de ‘simulación’ de las experiencias a las que se refieren (en la que se produce una activación de más o menos las mismas regiones del cerebro que se activan en la experiencia original). Este modelo permite una

⁵⁵ De hecho, Papineau llama la atención sobre el hecho de que a pesar de que ahora podemos explicar con física cuántica por qué el H₂O es incolora, inodora, insabora, etc., los científicos estaban convencidos de que agua era H₂O antes de que llegaran estas explicaciones, orientados por consideraciones semejantes al argumento causal.

explicación fisicalista alternativa de la adquisición de conceptos fenoménicos: ¿por qué sólo podemos adquirir un concepto fenoménico después de haber tenido previamente la experiencia correspondiente? Porque los seres humanos sólo pueden reactivar patrones de neuronas que previamente han sido activados por otras fuentes. Papineau concluye que los conceptos fenoménicos refieren no–descriptivamente⁵⁶ por medio de una especie de simulación.

Sin embargo, más allá de la teoría de la simulación que ofrece como respuesta a la idea de que existen descripciones asociadas *a priori* con los conceptos fenoménicos, Papineau piensa que no es esta idea la que explica la sensación de brecha. Lo que sucede realmente es que la mayoría de las personas no están preparadas para aceptar que los conceptos fenoménicos refieren a propiedades materiales, incluso aunque estén frente a los fuertes argumentos causales que muestran que así es, y siguen suponiendo de que se refieren a propiedades conscientes distintas de propiedades físicas; por lo tanto, es la persistencia de esta suposición la que hace que sea incomprensible que el fisicalismo.

¿Por qué la evidencia causal no es suficiente como en las identidades entre nombres propios? Según Papineau, este problema se relaciona con la naturaleza de los conceptos fenoménicos. Éstos parecen nombres propios por referir directamente, pero se diferencian por hacerlo por simulación, y es este segundo rasgo produce una poderosa ilusión de diferenciación mente–cerebro: al usar los conceptos fenoménicos se asocia el mismo carácter cualitativo de las experiencias a las que se refieren. Esta semejanza subjetiva entre el despliegue de los conceptos fenoménicos y las experiencias a las que se refieren, fácilmente puede confundirnos cuando consideramos la identidad, pues si nos enfocamos en un lado de la identidad, desplegamos nuestro concepto fenoménico y sentimos algo parecido; pero si nos enfocamos en el otro lado, desplegamos nuestro concepto neurofisiológicos y no sentimos nada parecido a la experiencia. Y por esto

⁵⁶ Vale la pena subrayar, a propósito de la teoría de la referencia de los conceptos fenoménicos por la que Papineau se inclina aquí, la observación que hace en el sentido de que la identidad fenoménico–neuronal encuentra mejores analogías en identidades entre nombres propios que entre identidades que involucran descripciones, pues los nombres propios refieren directamente y no están asociados con descripciones.

concluimos que ese lado deja por fuera la experiencia misma y se refiere sólo a los distintos correlatos físicos del dolor.

Cuando usamos nuestros conceptos fenoménicos imaginativamente, traemos a la mente un caso de la propiedad experiencial en la que estamos pensando. Cuando usamos conceptos no-fenoménicos no. Y esto nos lleva a pensar que no es posible que los conceptos no-fenoménicos puedan denotar las mismas propiedades experienciales que son seleccionadas por nuestros conceptos fenoménicos.

Sin embargo, subraya Papineau, esta conclusión es una falacia de tipo uso-mención (la denomina la *falacia antipática*): en cierto sentido los conceptos no-fenoménicos dejan por fuera las experiencias conscientes (no las *usan*, pero eso no quiere decir que no las *mencionen*). Los conceptos no-fenoménicos difieren de los fenoménicos en que no *usan* las experiencias a las que se refieren, aunque las denotan (y en este sentido dejan por fuera las experiencias); pero de esto no se sigue que los conceptos no fenoménicos difieran de los fenoménicos en cuanto a lo que *mencionan*, pues en este aspecto referencial no dejan por fuera ningún elemento de la experiencia ni siquiera su carácter subjetivo (pues este debe ser idéntico a algún rasgo cerebral).

2.5.4 Teniendo Más Cuidado con la Brecha

Sin duda, Papineau está en lo correcto al decir que no cabe pedir que se explique una identidad verdadera: una vez hemos establecido que la identidad es el caso, simplemente lo es, y no cabe preguntar por qué lo es. En efecto, lo que cabe preguntar es qué nos justificó para pensar que la identidad es el caso, e indiscutiblemente las consideraciones causales a las que apela Papineau son suficientemente poderosas para hacernos pensar que se trata de un caso de identidad. Pero algo que sigue siendo verdad es que en esta identidad no sucede algo que encontramos en otros casos familiares de identidad y que, según creo, es donde radica la existencia de una brecha explicativa (al menos como la he interpretado).

Específicamente, la asimetría que quiero señalar entre las identidades que nos son familiares (por ejemplo, Agua/H₂O), y la identidad fisicalista es la siguiente: pese a las consideraciones causales, en la segunda identidad (a diferencia de la primera) seguimos siendo incapaces de hacer una derivación del *explanans* (o de sus propiedades) a partir

de las propiedades del *explanandum*⁵⁷; seguimos sin entender por qué deberíamos esperar que, dadas ciertas propiedades de la entidad teórica que oficia como *explanandum*, se presente el *explanans* (o ciertas propiedades que nos son familiares en él).

Esta asimetría no se disipa con la afirmación de que los conceptos fenoménicos no refieren descriptivamente sino directamente por simulación, pues esta es la afirmación de que las dos clases de conceptos son lógicamente independientes; tampoco se resuelve con la afirmación (nageliana) de que imaginamos *simpáticamente* las experiencias conscientes e imaginamos *perceptualmente* los estados cerebrales, pues esta es otra forma de decir que las experiencias conscientes y los estados cerebrales hacen parte de dos modos de concebir diferentes (lo cual es causa de que haya una brecha explicativa); ni se resuelve diciendo que los dos términos son idénticos por la evidencia causal, pues esto no nos ayuda a comprender el hecho de que se produzca la mencionada asimetría (si bien aceptamos esta evidencia causal, la dificultad sigue surgiendo).

Como respuesta a lo anterior, sería de esperar una respuesta en la que se invoque la idea de que este es un tipo de exigencia aplicable únicamente a otro tipo de identidades, específicamente a aquellas en que los términos están asociados *a priori* con descripciones. Pienso que puede merecer ciertas reservas la idea de que los conceptos fenoménicos *no* están asociados *a priori* con descripciones (o que toda descripción es funcional, en oposición a fenoménica), pero incluso al conceder que esas sean exigencias para modelos de identidad diferentes de la identidad en cuestión, la afirmación de que la identidad psicofísica es asimétrica con respecto a las identidades que nos son familiares sigue intacta. Al menos como la he caracterizado, la brecha explicativa es la afirmación de que la identidad psicofísica es asimétrica en este sentido (con respecto a las que nos son familiares): hay algo que en un proceso de explicación *no se puede hacer* en este caso (independientemente de las razones por las cuales esto sea así).

⁵⁷ Recuérdense mis consideraciones sobre la brecha explicativa en el marco de la revisión que hago de la postura de Joseph Levine y David Chalmers, donde señalo *in extenso* en qué sentido estoy comprometido con la existencia de una brecha explicativa y por qué no se puede cerrar, además de consideraciones sobre qué se requeriría para cerrar la brecha y por qué algo como eso no puede obtenerse.

Si lo que he dicho es correcto, Papineau no logra disipar la brecha explicativa que he estado discutiendo. Ahora bien, en tanto enfoque fisicalistas hay algo que comparte la posición de Papineau con la que he venido describiendo: dentro de la posición de Papineau, el peligro de la brecha es desestimable en la medida de que no hay un brecha explicativa (pues no hay nada que explicar), dentro de mi posición el fisicalismo tampoco peligra ante la brecha explicativa, pero no porque no haya nada que explicar sino porque aunque la explicación sea imposible se trata de una imposibilidad epistemológica sin consecuencias a como parece ser el mundo en sí mismo. Mi planteamiento es que la brecha explicativa es más o menos inofensiva, pues aunque establece una barrera insalvable para las pretensiones explicativas de algún fisicalista, no afecta el hecho de que la identidad psicofísica sea el caso. He sostenido que la brecha no se puede cerrar porque lo que está en juego aquí es la relación entre dos modos de acceso epistémico inconmensurables, conceptualmente independientes, lo cual no significa que sean ontológicamente distintos (lo cual realmente merecería la preocupación del fisicalista).

2.5.5 Papineau y el Rechazo de las Propiedades Fenoménicas

Hasta cierto punto, la postura de Papineau resulta cercana a la de Loar, pues para los dos teóricos nuestros conceptos fenoménicos refieren de modo no-descriptivo (demostrativamente en el caso de Loar y por simulación en el caso de Papineau). La independencia conceptual con respecto a los conceptos de la ciencia física que se subraya con esto es lo que explica que no haya razones *a priori* para pensar que los referentes de ambas clase de conceptos serán los mismos.

Sin embargo, resulta obvio que Papineau no se pliega al enfoque de Loar sobre los conceptos fenoménicos (según el cual estos refieren propiedades fenoménicas), sino que lo critica.⁵⁸ Papineau parece pensar que la noción de ‘propiedad fenoménica’ contiene un

⁵⁸ “Some philosophers hold that any satisfactory account of the distinctive workings of our phenomenal concepts will have to mention phenomenal properties which are distinct from any material properties” (Papineau, 1998, p. 382); “This simulational model suggests an alternative physicalist account of the acquisition and subsequent deployment of phenomenal concepts, an account which makes no reference to any supposed acquaintance with distinctive phenomenal properties” (Ibíd.); “There is no reference in this story to any acquaintance with distinctive phenomenal properties” (Ibíd.); “Neither of these uses of phenomenal concepts demands

problema para el fisicalismo, en la medida que se trata de propiedades “distintas de cualesquiera propiedades materiales” (Papineau, 1998, p. 382); evidentemente no está considerando, como Loar, que dichas propiedades fenoménicas pueden ser idénticas a propiedades físicas. Pero si se omite este sesgo de Papineau, puede verse una convergencia entre su aproximación y la de Loar; si consideramos que las propiedades fenoménicas pueden ser idénticas a propiedades físicas (las propiedades fenoménicas y los patrones de activación neuronal que se producen en la simulación serían idénticos), la referencia demostrativa a propiedades fenoménicas (como “*esa experiencia*” o “*esa propiedad de esa experiencia*”) y la referencia por simulación de Papineau podrían resultar idénticas (ontológicamente e incluso cualitativamente). Desde luego, esta no es más que una sugerencia que requiere un análisis más detenido, pero al menos a primera vista parece atractiva.

Como consideración final. Una de las ideas centrales de Papineau para negar la existencia de una brecha explicativa es que los conceptos fenoménicos no tienen descripciones asociadas a priori. Con todo, vale la pena resaltar que los argumentos que presenta Papineau están dirigidos a mostrar que los conceptos fenoménicos no están asociados a priori con descripciones *funcionales*. ¿Son todas las descripciones funcionales? La existencia de propiedades fenoménicas nos permite pensar que no necesariamente es así: dado que cierta experiencia consciente tiene la propiedad de sentirse de tal y tal manera, podríamos referirnos a ella mediante la descripción definida “la experiencia que se siente de tal y tal manera” (que indica que dicha experiencia es la que tiene tal propiedad). Al menos no es obvio que esta descripción sea funcional. Así, alguien podría preguntarse cómo es que un estado neurológico, en la medida en que es idéntico a la experiencia consciente en cuestión, satisface dicha descripción; tendríamos una brecha explicativa.

3. CONCLUSIONES

Hasta ahora me propuse hacer una revisión crítica de algunos de los principales argumentos que fundaron el debate alrededor de estatus ontológico de la experiencia consciente, concentrándome en la discusión sobre el fisicalismo acerca de la conciencia fenoménica. Deseo ahora recapitular los resultados que creo que se recogen de esta revisión, hacia dónde me parece que apuntan y qué es lo que considero que establecen.

3.1 Sobre el Desafío de Kripke

La dificultad con la que Kripke enfrenta al fisicalista puede ponerse en los siguientes términos: ¿a qué modelo podría recurrir para disipar la contingencia de la identidad psicofísica?

He sostenido que no se ha probado que el modelo de identificación por medio de una propiedad contingente asociada sea el único que pueda disipar este tipo de contingencia. He planteado que puede intentarse una repuesta a este desafío que opere con una distinción entre lo epistémico y lo ontológico (admitida por el mismo Kripke), según la cual el que concibamos una cierta situación no hace que sea posible. En efecto, esta es la postura que plantea Levine; en esta postura, la ilusión de contingencia de la identidad psicofísica no equivaldría al hecho de que creamos que podemos concebir una situación contrafáctica que de hecho no podemos concebir, sino a que –pese a que podamos concebirla– su concebibilidad no significan que pueda ser posible.

Si bien Kripke aseveraría que *no es concebible que un enunciado de identidad verdadero pudiera ser falso*, Levine argumenta que *se puede concebir que un enunciado de identidad verdadero sea falso, porque lo concebible no es una guía para lo que es [metafísicamente] posible*. Siguiendo a Levine, esta diferencia entre el punto de vista suyo y el de Kripke se debe a que aquél presupone que nuestro entendimiento es un indicador infalible de lo que es posible.

Si nos orientamos por la invitación de distinguir lo epistemológico de lo ontológico, no parece haber ninguna razón para rechazar la concebibilidad: resulta razonable que lo que se puede considerar una situación contrafáctica posible no lo sea.

Considero pertinente cerrar con la siguiente anotación. La pretensión de estas consideraciones no es hacer una refutación de Kripke, ni acusarlo de meramente caer en un error psicológico. Más bien se trata de señalar algunas tensiones inherentes a su posición, por las cuales no puede resultar concluyente en contra del fisicalismo. Lo que he tratado de mostrar es que varios elementos del enfoque de Kripke no son obvios e indiscutibles; queda, por tanto, un rango de discusión que es interesante abrir. Es aquí donde una aproximación tipo Levine parece una buena alternativa: interpretar de una forma mucho más débil el alcance de las premisas de Kripke (epistemológicamente y no metafísicamente).

3.2 Sobre los argumentos de Nagel

Señalé que el principal aporte que hacen las reflexiones de Nagel sobre la relación entre las experiencias conscientes y los estados físicos es la observación de que el carácter subjetivo de la experiencia no es capturado por ninguno de los análisis reductivos de lo mental. Esto se ve demostrado en el hecho de que todos los análisis reductivos son lógicamente compatibles con la ausencia de subjetividad. Lo anterior se debe a que el análisis fisicalista (o funcionalista) de lo mental es conceptualmente independientes del un análisis subjetivo de lo mental (por esta independencia no pueden capturarlo, pues no lo implican). De hecho, estos análisis reductivos suprimen el carácter subjetivo por su pretensión de eliminar el punto de vista del sujeto; por esto es imposible entonces reducir la experiencia consciente a la terminología de la ciencia física, pues en el proceso no quedaría nada de la experiencia consciente misma.

Pero la afirmación clave de Nagel es que esta incapacidad explicativa no conlleva la negación del fisicalismo, hay *hechos* que no consisten en la verdad de proposiciones expresables en un lenguaje humano. Aunque no podamos explicar un hecho, aunque no entendamos completamente cómo puede ser verdad, no significa que no sea verdad (más aún, podemos tener *evidencia* de algo que no podemos entender completamente).

Con todo, la consecuencia de la irreductibilidad del carácter subjetivo de la experiencia no es menor: se trata de la existencia de una brecha entre dos modos de concebir.

3.3 El Argumento del Conocimiento

La lectura que mantengo sobre el Argumento del Conocimiento es que, interpretado en el primer sentido de información física (“información explícitamente física”), el fisicalista no tendría mayores reparos para aceptarlo, pues no constituiría una amenaza insalvable contra su postura (en la medida que no se seguiría la falsedad del fisicalismo, sino que sería más bien un punto de tipo nageliano). Si interpretamos el argumento del conocimiento en el segundo sentido de ‘información física’, el asunto es diferente.

Propuse desistir del intento de minar la segunda premisa de la ‘versión fuerte’ (derivada del segundo sentido de información física: información ontológicamente física) y volver a la primera premisa. Mi sugerencia fue que si el conocimiento ontológicamente físico pudiera dividirse en proposicional y por aprehensión (by acquaintance), podríamos decir más bien que dentro de la habitación en blanco y negro se puede obtener todo el conocimiento ontológicamente físico *proposicional* pero no todo el conocimiento ontológicamente físico *por aprehensión*. Sin argumentos para descartar esta posibilidad – más aún, sin tenerla en cuenta– no parece que podamos dar el paso a afirmar que *todo* el conocimiento ontológicamente físico puede ser adquirido dentro de la habitación de Mary.

Señalé además que, a la vez que se aferra a la intuición de que dentro de la habitación no puede adquirirse cierto tipo de conocimiento, al afirmar que *todo el conocimiento ontológicamente físico* puede adquirirse dentro de la habitación –ignorando la sugerencia que hice–, el defensor de la versión fuerte presupone que cualquier conocimiento que no puede ser adquirido dentro de ella es ontológicamente no-físico; se *presupone* que el fisicalismo es falso, no se demuestra. En la medida en que estamos dispuestos a admitir que Mary aprenderá algo nuevo al dejar su habitación, no podemos simplemente admitir que ella ha adquirido todo el conocimiento ontológicamente físico cuando estaba dentro de la habitación. La primera premisa del argumento en la versión *b* podría ser falsa, caso en el cual el argumento no se seguiría: Mary podría no tener todo el conocimiento ontológicamente físico.

3.4 La Brecha Explicativa

La versión que he defendido de la brecha explicativa puede presentarse de una manera bastante sumaria. Si algo es explicado cuando ‘no queda nada por fuera’ (en el sentido de que se podía hacer una derivación completa del *explanandum* a partir del *explanans*), entonces la brecha explicativa es cuestión epistemológica *de principio* de que no se pueda hacer una derivación de los hechos subjetivos a partir de hechos neurofisiológicos. No hay *nada* en la activación de las fibras C que haga que se ajuste, mejor que otro estado neuronal, a las propiedades fenoménicas del dolor. No existe ningún vínculo *explicativo* entre los hechos fenoménicos y los hechos físicos (el discurso sobre estados cerebrales no *implica* lógicamente nada acerca del carácter cualitativo de una experiencia; esto es lo que significa que sean conceptualmente independientes los unos de los otros); en lugar de este vínculo explicativo, tenemos una *brecha explicativa*.⁵⁹

Pero la brecha explicativa resulta un problema serio para el fisicalismo sólo si este incluye la tesis de que, en la medida en que todo es físico, *debe ser posible* hacer una derivación de cualquier hecho a partir de los hechos descritos en términos de la ciencia física (*‘fisicalismo epistemológico’*). El *‘fisicalismo ontológico’* (aquel que se escinde de cualquier posición epistemológica) no ha sido puesto en peligro.

Me interesa resaltar, con miras a la posición que delinearé más adelante, que Levine interpreta el caso de Mary como un ejemplo de que puede haber más de una manera de tener acceso epistémico a un mismo hecho, y que no se puede pasar siempre de múltiples modos de acceso a múltiples hechos accedidos. Levine (1993) termina su discusión de la ‘lectura metafísica’ señalando que lo único que hacen tanto el argumento de Kripke como el de Jackson es poner en evidencia la división epistemológica entre dos modos de acceso a lo que todo lo demás⁶⁰ indica que es el mismo fenómeno. De un lado, tenemos ciertas descripciones físico (extensionales, de tercera persona) de ciertos estados de ciertos sujetos. De otro lado, tenemos descripciones derivadas de la experiencia de primera mano de dichos estados. Pero no tenemos razones para pensar

⁵⁹ Esto no significa, desde luego, que la única conexión que puede encontrarse entre ambos ítems sea *explicativa*. He señalado cómo podrían encontrarse buenas razones para pensar que existe un vínculo a partir de consideraciones causales.

⁶⁰ Es decir, múltiples consideraciones causales y metodológicas.

que los argumentos en cuestión muestren que estas dos clases de descripciones no se refieren a los mismos hechos, por lo cual tampoco tenemos razones para concluir que el fisicalismo ‘deja por fuera’ algo en el sentido de que no se refiera de algún modo a tales hechos.

3.5 El Hard Problem y Brecha Explicativa

He aseverado que el Problema Difícil se deriva de la existencia de la brecha explicativa. La pregunta C no se puede responder porque no se pueda hacer una derivación de los hechos subjetivos de la experiencia a partir de hechos neurofisiológicos.

Si “una teoría completa de la conciencia debe edificar un puente explicativo” entre lo físico y la experiencia, entonces no puede existir una teoría satisfactoria de la conciencia. Pues ¿qué se requeriría para cerrar ese brecha? Se requeriría una aproximación que nos permitiera pasar del modo de concebir ‘de tercera persona’ al modo de concebir ‘de primera persona’; pero sucede que este conocimiento involucraría exigencias contradictorias: suprimir el punto de vista del sujeto y, al mismo tiempo, incluirlo.

3.6 La Premisa Semántica y la Brecha Explicativa

He tratado de examinar uno de los presupuestos fundamentales que contiene el argumento de la *Premisa Semántica*, a saber, la idea de que si dos conceptos capturan la esencia o designan rígidamente una misma propiedad, deberíamos poderlo saber *a priori*.

Para empezar, se está interpretando la expresión ‘designar rígidamente’ como ‘es conceptualmente inter-derivable con algún predicado teórico’; pero no parece que la designación rígida implique que dos conceptos deban ser inter-derivables *a priori*.

He observado que si los conceptos ‘Agua’ y ‘Calor’ son designadores rígidos, y los conceptos teóricos también lo son, se sigue que *ambos* capturan la esencia de su referente; pero esto es algo que claramente no sabemos *a priori* (pues descubrimos que el agua es H₂O y que el calor es la tasa de movimiento molecular). Así que si el argumento de la *Premisa Semántica* fuera correcto, identidades como agua=H₂O serían falsas si sólo pueden ser conocidas *a posteriori* o *a priori* si han de ser verdaderas.

Con respecto a la brecha explicativa, he mostrado mi adhesión a la orientación de Loar según la cual la brecha se genera porque suponemos que los enunciados de identidad verdaderos *deberían* ser explicativos, pero los enunciados de identidad que conectan conceptos fenoménicos y conceptos físicos/funcionales no lo son. No obstante, podríamos encontrar enunciados de identidad que pueden no ser explicativos (por su independencia conceptual) y ser verdaderos, así que lo que tenemos que justificar más bien es la intuición de que todo enunciado de identidad debería ser explicativo.

3.7 Papineau y su negación de la Brecha Explicativa

Luego de examinar la argumentación de Papineau en contra de la brecha explicativa, he sostenido que –pese a las consideraciones causales– a diferencia de lo que sucede en identidades como “agua=H₂O”, en identidades como “dolor=activación de las fibras C” seguimos siendo incapaces de hacer una derivación del *explanans* a partir de las propiedades del *explanandum*; seguimos sin entender por qué deberíamos esperar que, dadas ciertas propiedades de la entidad teórica que oficia como *explanandum*, se presente el *explanans* (o ciertas propiedades que nos son familiares en él).

Argumenté que la asimetría entre los dos tipos de identidad no se disipa con la afirmación de que los conceptos fenoménicos no refieren descriptivamente sino directamente por simulación, ni con la afirmación de que imaginamos *simpácticamente* las experiencias conscientes e imaginamos *perceptualmente* los estados cerebrales, ni diciendo que los dos términos son idénticos por la evidencia causal.

También argüí que si se invoca la idea de que este es un tipo de exigencia aplicable únicamente a otro tipo de identidades, pero incluso si concediendo esto, la afirmación de que la identidad psicofísica es asimétrica con respecto a las identidades que nos son familiares sigue intacta. Y, tal como la he caracterizado, la brecha explicativa consiste sencillamente en eso: que las identidades tipo “dolor=activación de las fibras C” son asimétricas con respecto a identidades tipo “agua=H₂O”.

3.8 Sobre el Fisicalismo

Desde la introducción de esta investigación me he referido al fisicalismo como la tesis de que todo lo que existe es físico. Se trata de una tesis eminentemente ontológica, sobre lo

que hay en el mundo, de qué está hecho, cuál es su mobiliario; específicamente, es una forma de monismo ontológico. Se requiere, no obstante hacer algunas precisiones sobre esta formulación –deliberadamente vaga– de lo que es el fisicalismo.

Muchas veces se asume que aceptar el fisicalismo implica aceptar toda una serie de posturas epistemológicas además de la pura tesis ontológica. Por ejemplo, se piensa que aceptar el fisicalismo implica aceptar que el método de investigación de las ciencias físicas proporcionará una comprensión de *todo* en el mundo. También se cree que el fisicalismo incluye cierto ‘optimismo epistémico’, esto es, la creencia de que el método de investigación de las ciencias físicas puede ser usado para alcanzar una explicación total o final de todo lo que hay en el universo. Esta última creencia se asocia con la idea de que todas las formas de conocimiento deberán unificarse en una única ciencia (la física, o algo muy semejante a la física). Otra idea que se asocia con el fisicalismo es que el genuino conocimiento del mundo no involucra ninguna clase de referencia a puntos de vista particulares o experiencias (es decir, que el fisicalismo implica objetivismo y supresión total de lo subjetivo).

Pero el fisicalismo puede separarse por igual de todas estas ideas por una sola razón: porque es una tesis metafísica u ontológica y no una tesis metodológica o epistemológica; “*físico*” aquí no significa simplemente “*describible por la terminología de las ciencias naturales*”, sino que tiene un significado más amplio. La tesis ontológica del fisicalismo no implica ninguna posición epistemológica o metodológica sobre cómo debería proceder la ciencia natural, o sobre la eliminación de las aproximaciones epistémicas que incluyen experiencias o puntos de vista.

Existen, por otra parte, un conjunto de sólidas razones para adherir el fisicalismo. De entrada, tenemos abundante evidencia de que ciertas modificaciones en la actividad cerebral pueden alterar de manera sorprendentemente específica la experiencia consciente de un individuo, evidencia que parece difícil acomodar si no aceptamos que las experiencias conscientes son eventos físicos. Esta consideración por sí misma no prueba el fisicalismo, a menos que añadimos otro argumento para apoyar el fisicalismo, que se basa en la Navaja de Ockham (“no multiplicar las entidades innecesariamente”). En este sentido, pensar que la conciencia fenoménica es un tipo de rasgo físico del

cerebro sigue siendo una explicación más parsimoniosa que cualquier otra alternativa a la mano.

Por lo demás, la fuerza del fisicalismo parece radicar también en su alcance o poder explicativo. Se trata de una concepción metafísica compatible con todo lo que sabemos, que ubica en el mismo marco un sinnúmero de fenómenos, desde los más simples de la física hasta los más complejos de la química, la biología, la psiquiatría, etc. En efecto, pareciera que el único aspecto de la realidad que la faltara al fisicalismo por acomodar dentro de su enfoque fuera la conciencia fenoménica.

Sin embargo, la razón más sólida para apoyar el fisicalismo es la que se ha conocido como el 'argumento causal' (mi versión es ligeramente diferente de la de Papineau): todo evento físico debe tener una causa física y sus efectos son efectos físicos (de acuerdo con el principio del cerramiento causal del mundo físico); los estados mentales son causados y pueden causar eventos físicos. Por tanto, en la medida que causan eventos físicos sin sobredeterminación causal, los estados mentales deben ser eventos físicos.

Dicho de otro modo, sabemos que nuestros estados neurofisiológicos son causas suficientes de nuestros movimientos corporales: la secuencia causal completa de cualquier movimiento corporal empieza con una serie de activaciones neuronales que envían un impulso eléctrico a ciertos nervios y éstos lo reproducen hasta llegar a determinados paquetes musculares, etc., etc. Pero, por otro lado, también pensamos que nuestras experiencias conscientes causan nuestros movimientos corporales: decimos que la causa de que levantemos el brazo es que nos asustamos o que nos dolió. Ahora, la sobredeterminación causal no parece el modelo adecuado para la causación mental (recuérdese lo dicho por Papineau). Así pues, parece que podríamos zanjar este asunto concluyendo que las experiencias conscientes y ciertos estados neurofisiológicos – ambos causas suficientes de los movimientos corporales– son el mismo fenómeno.

3.9 Discreta delineación de una auténtica salida para fisicalista: La Teoría *Perspectival* de la Conciencia (TPC)

Una manera de plantear el problema que se le presenta al enfoque fisicalista es el siguiente: la descripción adecuada de las experiencias –en tanto fenómenos de primera persona– es desde el punto de vista de la primera persona; pero a pesar de esto, el fisicalista debe mantener que siguen siendo fenómenos físicos, y debe buscar la forma de explicar que este hecho epistemológico sobre la experiencia consciente no tiene consecuencias ontológicas indeseables. Para lograr esta tarea, una teoría fisicalista debe involucrar un enfoque que, a falta de una mejor palabra, llamaremos metafóricamente “*perspectival*”.

Con la expresión “enfoque *perspectival*” me refiero a un tipo de análisis dentro de cual se conciba lo fenoménico (las experiencias conscientes) como una ‘*perspectiva*’, ‘modo de presentación’, ‘modo de darse’ o ‘*aspecto*’ de ciertos estados físicos. Dicho de otro modo, las experiencias conscientes serían la manera que tienen de presentarse (la manera de dársele) ciertos estados físicos al sujeto que los sufre. Tendríamos así dos maneras de darse de la misma cosa: una ante el sujeto de dichos estados (como experiencias conscientes) y otra ante el observador de dichos estados (como actividad eléctrica en ciertas regiones de una materia gris y gelatinosa). Las experiencias conscientes serían estados neurofisiológicos ‘*vistos*’ desde el punto de vista de primera persona, es decir, tal como se le dan al sujeto que los posee. El observador externo, por su parte, tendría una perspectiva ‘de tercera persona’ de estos estados físicos.

Otra manera de llegar a la TPC es a través del argumento causal. Si es verdad que lo físico y lo fenoménico han de ser una misma realidad, sólo hay una posibilidad: que cada uno es dicha realidad desde ‘*perspectivas*’ epistémicas distintas.

La TPC es una defensa de la teoría de la identidad pues de acuerdo con ella las experiencias conscientes serían ciertas propiedades y procesos físicos. El asunto se reduciría a que esos procesos físicos tienen diferentes modos de darse (son vividos de diferentes modos según seamos o no sus sujetos): aunque el neurocientífico y yo podemos estar enfocándonos en la misma cosa cuando él ve mi cerebro, sólo yo soy su

sujeto, y eso hace que sólo yo pueda vivíroslos como experiencias conscientes (y que el neurocientífico sólo pueda percibirlos como activaciones neuronales). A diferencia de otras versiones anteriores de la teoría de la identidad, no es que las experiencias conscientes y sus rasgos cualitativos *sean realmente* ciertos procesos físicos, sino que desde una perspectiva lo son y desde otra perspectiva son experiencias conscientes.

La experiencia consciente, entonces, involucra estar en ciertos estados físicos, que son *vividos de cierta manera* por los organismos que los sufren. Y aunque hay un sentido en el que el neurocientífico puede observar mis experiencias conscientes sólo tiene acceso de tercera persona a ellas (pues no es su sujeto).

Tenemos dos vías epistémicas para acceder a los estados físicos en cuestión, que podríamos llamar la '*perspectiva de primera persona*' y la '*perspectiva de tercera persona*'; desde la primera, concebimos dichos estados como experiencias conscientes, y desde la segunda los concebimos como estados cerebrales. De cada una de estas vías epistémicas se derivan dos clases de conceptos con los cuales nos referimos a estos estados físicos: los conceptos fenoménicos y los conceptos de la ciencia física. Los conceptos de la ciencia física son aquellos que refieren a los estados en cuestión desde una perspectiva de tercera persona; capturan cosas como los roles causales y la constitución estructural de dichos estados. Por otra parte, los conceptos fenoménicos son aquellos que usamos para referirnos a las experiencias conscientes por el carácter cualitativo subjetivo que sentimos cuando las tenemos.

La TPC señala que la relación experiencia consciente/ciertos rasgos cerebrales tiene la forma de una identidad empírica *a posteriori* y, que existe una brecha explicativa entre los dos ítems de la identidad (debida a la independencia lógica de los conceptos de la ciencia física y los conceptos fenoménicos). Adicionalmente, es consistente con los resultados que ha arrojado la presente investigación, como puede verse en los numerales 1 al 8 de esta tercera parte.

3.9.1 La TPC y el Dualismo Conceptual

El núcleo de la TPC consiste en reconocer que la experiencia consciente y la descripción científica de su correlato neuronal son dos formas distintas de acceso epistémico al mismo hecho. Así, la experiencia consciente no ofrece conocimiento de nuevos hechos

que permanecieran ignorados desde una perspectiva fisicalista del mundo, sino sólo otra manera de conocer los *mismos* hechos descritos en términos de la ciencia física. Esto guarda mucha consonancia con una de las respuestas al Argumento del Conocimiento, el argumento del '*viejo hecho/nueva manera*' (Levine 1993, Horgan 1984, Churchland 1985); en esta posición, Mary efectivamente adquiere algo al salir de la habitación, pero no es conocimiento de un nuevo hecho sino más bien una nueva manera de conocer un hecho que ya conocía antes de otra manera.

A partir de la estrategia del '*viejo hecho/nueva manera*' se han desarrollado una serie de aproximaciones a las que se ha denominado 'Dualismo Conceptual'. Papineau, por ejemplo, considera que el argumento de Jackson no logra concluir la existencia de propiedades fenoménicas irreductibles a propiedades físicas, pero sí logra demostrar que existen *conceptos fenoménicos* (cf. Papineau 2002, 50) que son distintos e irreductibles a *conceptos materiales*. Papineau no cree que después de que Mary sale de su cuarto conozca propiedades *no físicas* sobre el mundo; piensa que Mary sí aprende algo nuevo al salir, pero no cree que este nuevo aprendizaje sea sobre propiedades no físicas. Lo que aprende Mary cuando ve una manzana, dice Papineau, es un nuevo *concepto* gracias al nuevo acceso epistémico que adquiere al salir de la habitación. La propiedad a la que Mary se refiere a través de este nuevo concepto es una propiedad física que, también, puede ser perfectamente descrita a través de conceptos físicos. Tenemos entonces un dualismo en el ámbito de los conceptos y un monismo en el ámbito ontológico.

La TPC se sitúa muy cerca del dualismo conceptual, pero contiene diferencias significativas; la más significativa es que en el dualismo conceptual, la subjetividad de la experiencia se deriva de la manera en que están formados los conceptos fenoménicos, mientras que en la TPC los conceptos fenoménicos tendrían una constitución tal (que los hace irreductibles a conceptos de la ciencia física) derivada de la intrínseca subjetividad de la experiencia consciente.⁶¹

⁶¹ Esta diferencia salta a la vista por lo siguiente: mientras que es usual encontrar entre los dualistas conceptuales algunas observaciones en contra de la aproximación de Nagel, en esta investigación la he admitido ampliamente.

3.9.2 La TPC y la Brecha Explicativa

Una consecuencia interesante de la posición que planteo es que aunque las experiencias conscientes sean entidades físicas, lo más que podemos esperar es una mera correlación entre las descripciones de primera persona y las descripciones de tercera persona. Esta es una brecha explicativa epistemológica insalvable, sin consecuencias ontológicas, producida por la independencia conceptual entre dos 'modos de concebir' o modos de darse de una misma realidad. Es esta independencia lógica la que determina que la verdad de la identidad psicofísica sólo pueda conocerse *a posteriori*, y la que determina que sean concebibles ciertas situaciones que falsearían la identidad, generando por lo tanto la brecha explicativa.

3.10 Obstáculos Potenciales para la TPC

Ya me he referido a algunas de las objeciones que pueden anteponerse a este tipo de versiones del fisicalismo, pero merecen ser revisitadas con algo de detenimiento.

3.10.1 El 'Argumento del Análisis Conceptual'

La TPC, es una versión de la teoría de la identidad. De acuerdo con ella, los referentes de los conceptos fenoménicos son los mismos que los referentes de los conceptos físicos; lo que nos hace pensar lo contrario es que, en la medida que cada vía epistémica tiene su propio sistema de conceptos, suponemos que sus referentes deben ser diferentes. De este modo, admite que algún enunciado de la forma "el estado cerebral *x* es idéntico al estado fenoménico *y*" es verdadero.

Podemos deducir fácilmente que el dualismo conceptual resulta vulnerable a críticas derivadas de su compromiso con este tipo de enunciados de identidad. Frank Jackson (1993) y David Chalmers (1996) han concentrado su crítica en la idea de que admitir este tipo de enunciados implica un compromiso con la característica principal de los enunciados de identidad: su necesidad. De acuerdo con Jackson y Chalmers, incluso cuando tratamos con afirmaciones de necesidad que son *a posteriori* debe haber un criterio que nos permita justificar dicha afirmación de necesidad (una suerte de *test de necesidad*). Tomando ejemplos típicos de identidades conocidas (como "agua=H₂O", "calor =movimiento de las partículas"), afirman que en ellos se puede hacer lo que llaman

“análisis conceptual” y deducen que en esto debe consistir el criterio que justifica la afirmación de necesidad.

Tomemos, por ejemplo, la siguiente afirmación: la tierra está 60% cubierta de agua. Si es necesario que $\text{agua} = \text{H}_2\text{O}$, deberíamos ser capaces de mostrar que el enunciado ‘la tierra está 60% cubierta de agua’ se deriva *a priori* de ciertas verdades microfísicas junto con verdades *a priori*, mediante un ‘análisis conceptual’.

En líneas generales, la idea es que es posible hacer la siguiente derivación con términos como ‘agua’ o ‘calor’, pero no con términos que expresan conceptos fenoménicos:

(i) Agua= el líquido incoloro, inodoro, etc. [Presunta definición analítica, verdad *a priori*]

(ii) H_2O = el líquido incoloro, inodoro, etc. [Verdad microfísica]

(iii) la tierra está 60% cubierta de H_2O . [Verdad microfísica]

Por tanto, dado que por (i) y (ii) $\text{agua} = \text{H}_2\text{O}$

(d) la tierra está 60% cubierta de agua.

La imposibilidad de realizar esta derivación en el caso de los conceptos fenoménicos se relaciona directamente con el hecho de que no puede haber definiciones en términos físicos o funcionales de las expresiones que usamos para describir nuestra experiencia fenoménica. En otras palabras, no tenemos nada equivalente a (i), por lo que es imposible hacer la derivación.

Así las cosas, resulta que las identidades entre expresiones fenoménicas y expresiones físicas no pasan el ‘test de necesidad’. Dado que el análisis conceptual que puede llevarse a cabo con términos como ‘agua’ y ‘calor’ (y que se requiere para justificar la afirmación de *necesidad*) no puede realizarse con términos que expresan conceptos fenoménicos, podemos concluir que el fisicalismo es falso: no hay una conexión metafísicamente necesaria entre la conciencia fenoménica y lo físico. Si estas identidades no son necesarias,⁶² como mostró Kripke, no pueden ser verdaderas, y

⁶² Kripke no dice realmente que todos los enunciados verdaderos de identidad sean necesarios, sino que lo son aquellos que contengan designadores rígidos; justamente lo que se supone en la discusión es que los conceptos fenoménicos y los conceptos físicos son designadores rígidos.

ninguna versión de la teoría de la identidad podrá ser correcta –lo cual incluye, por supuesto, al dualismo conceptual–.

3.10.2 La ‘*Premisa Semántica*’

La segunda crítica se enfoca en otro rasgo de los enunciados de identidad empíricos: su carácter *a posteriori*. De acuerdo con esta objeción, la única manera en la que se podría explicar el estatus *a posteriori* de un enunciado de identidad es que uno de los términos capture su referente contingentemente, ya que si ambos términos capturaran su esencia deberíamos saber *a priori* que los dos términos convergen en el mismo referente. Dicho de otro modo, una identidad *a posteriori* que vincula dos términos conceptualmente independientes es verdad sólo si al menos uno de ellos selecciona contingentemente su referente. Pero ni los conceptos fenoménicos ni los términos científicos capturan sus referentes de manera contingente. Por lo tanto, en la medida en que no podemos saber *a priori* que convergen en el mismo referente, y admitimos que los dos tipos de conceptos capturan la esencia de sus referentes, la identidad no puede ser cierta.

3.11 Posibles Alternativas

Ya insinué cuáles podrían ser las alternativas del fiscalista para responder a estos dos argumentos. Con respecto a la *Premisa Semántica*, véase la sección 7 de esta tercera parte.

Con respecto al *Argumento del Análisis Conceptual*, en la sección 7.3.1 de la segunda parte observé que se puede responder a la acusación de que si una identidad como dolor=activación de las fibras C fuera asimétrica con identidades tipo agua= H_2O no pasaría el ‘test’ propuesto por Chalmers y Jackson para justificar la afirmación de necesidad. La idea de que ese sea un test válido de necesidad parece problematizable: podría afirmarse que es una idea que pone un asunto de ontología (la necesidad) en términos de un asunto epistemológico (la derivabilidad *a priori*) a través de la noción de justificación.

3.12 Concebibilidad y Posibilidad⁶³

La discusión sobre concebibilidad y posibilidad es quizás una de las regiones más espinosas de la discusión que he seguido durante toda esta investigación. Las preguntas que surgen y las respuestas que se ofrecen son tremendamente intrincadas y controvertibles: ¿a qué llamamos “concebible”?, ¿a qué llamamos “posible”?, ¿es lo concebible una guía confiable para lo posible?, ¿qué tipos de posibilidad hay?, ¿cómo se relacionan? Se trata de un tema que evidentemente es relevante para la discusión que he examinado, pero cuyo tratamiento exhaustivo y riguroso desborda por mucho los propósitos de mi investigación presente (involucra tomar posición con respecto a problemas de lógica modal, de semántica de mundos posibles e incluso de teorías de la referencia y el significado). En consecuencia, me limitaré a las consideraciones siguientes.

No parece haber una caracterización no problemática de concebibilidad. Algunas veces se tiende a asociar lo concebible con las capacidades cognitivas del individuo, por lo cual se resulta interpretando “concebible” como “imaginable”. Se ataca entonces la noción de concebibilidad a través de la noción de imaginabilidad señalando que la imaginabilidad es completamente relativa a los sujetos, a sus posiciones epistémicas y a sus capacidades cognitivas. Sin embargo, al parecer existe una noción mucho más robusta de concebibilidad, que se encuentra mucho más relacionada con restricciones de tipo lógico, más que epistémico.

¿Qué significa en esta segunda lectura que algo sea concebible? Lo que podemos encontrar en Descartes es que la idea de que es metafísicamente posible todo aquello que no involucre contradicción: “no es dudoso que Dios pueda producir todas las cosas que soy capaz de concebir con distinción; y nunca he juzgado que le fuera imposible hacer una cosa, a no ser que ésta repugnase por completo a una concepción distinta” (Descartes, 1977, sexta meditación) y “no debo en modo alguno dudar acerca de la verdad de esas cosas, si, tras recurrir a todos mis sentidos, a mi memoria y a mi

⁶³ Estoy en deuda con el profesor Jaime Ramos, quien me aportó valiosas consideraciones sobre este tema particular, las cuales me permitieron tener algo de claridad en medio de este asunto tan farragoso.

entendimiento para examinarlas, ninguna de esas facultades me dice nada que repugne a las demás” (Ibíd.). Así la concebibilidad resulta dependiendo únicamente del respeto al principio de no-contradicción, separándose de las capacidades cognitivas de los individuos (de la imaginabilidad).⁶⁴ “Concebible” aquí sería intercambiable con “no-contradictorio”.

Las quejas que surgen respecto a esta noción de concebibilidad es que resulta excesivamente amplia, casi trivial; como señala Chalmers: “(...) *que el mundo fuera creado hace cinco minutos: tal vez es lógicamente coherente, pero no es plausible*”.

Una noción alternativa de concebibilidad es aquella según la cual lo concebible es lo compatible con todo lo que sabemos. Dado un cierto sistema de conocimiento, ciertas cosas resultarían concebibles y otras inconcebibles; la concebibilidad sería relativa a uno u otro de estos sistemas. Pero claramente estos sistemas cambian con la historia: lo que es concebible para una época o cultura no lo es para otra. A través del tiempo se puede producir la modificación de algunos conceptos, lo cual conllevará a que en algunos momentos ciertos conceptos sean incompatibles con otros y en otras ocasiones no lo sean.⁶⁵ Así, según todo lo que sabemos, parecería concebible que un sujeto tenga dolor sin activación de las fibras C, pero no sin *alguna* activación. Si se alegara en contra de

⁶⁴ “Y para manifestar esto con mayor claridad, advertiré primero la diferencia que hay entre la imaginación y la pura intelección o concepción. Por ejemplo: cuando imagino un triángulo, no lo entiendo sólo como figura compuesta de tres líneas, sino que, además, considero esas tres líneas como presentes en mí, en virtud de la fuerza interior de mi espíritu: y a esto, propiamente, llamo «imaginar». Si quiero pensar en un quiliógono, entiendo que es una figura de mil lados tan fácilmente como entiendo que un triángulo es una figura que consta de tres; pero no puedo imaginar los mil lados de un quiliógono como hago con los tres del triángulo, ni, por decirlo así, contemplarlos como presentes con los ojos de mi espíritu. Y si bien, siguiendo el hábito que tengo de usar siempre de mi imaginación, cuando pienso en las cosas corpóreas, es cierto que al concebir un quiliógono me represento confusamente cierta figura, es sin embargo evidente que dicha figura no es un quiliógono, puesto que en nada difiere de la que me representaría si pensase en un miriógono, o en cualquier otra figura de muchos lados, y de nada sirve para descubrir las propiedades por las que el quiliógono difiere de los demás polígonos. Mas si se trata de un pentágono, es bien cierto que puedo entender su figura, como la de un quiliógono, sin recurrir a la imaginación; pero también puedo imaginarla aplicando la fuerza de mi espíritu a sus cinco lados, y a un tiempo al espacio o área que encierran. Así conozco claramente que necesito, para imaginar, una peculiar tensión del ánimo, de la que no hago uso para entender o concebir; y esa peculiar tensión del ánimo muestra claramente la diferencia entre la imaginación y la pura intelección o concepción”. (Descartes, 1977, sexta meditación)

⁶⁵ Ahora, si hemos de admitir un mundo que existe independiente de que lo pensemos, no se seguiría que de esta noción de concebibilidad que lo concebible determine lo que es posible.

esto que es inteligible la idea de un ser inmaterial con dolor, no tenemos la carga de la prueba.

3.13 Perspectivas de Investigación Futura

Se ha dicho por mucho tiempo que al fisicalismo se le quedan por fuera ciertas entidades a la hora de explicarlas como físicas, que no puede explicar que sean físicas: estas entidades no podrían ser descritas en el lenguaje del fisicalista, no entrarían en la descripción fisicalista del mundo. Hace un momento insistí en que “*físico*” no significaba simplemente “*descriptible en términos de la ciencia física*”, y creo que cuando se usa la expresión “lenguaje del fisicalista” se está pensando en el lenguaje de las ciencias físicas.

Pero si “*físico*” no significa simplemente “*descriptible en términos de la ciencia física*” ¿qué significa? Mi sospecha es que aquí habrá que volver a la tradicional expresión “materialismo”, para no generar la confusión entre esta tesis ontológica y el aparataje de la ciencia física, para evitar asociarlos. Este tipo de materialismo sostendría entonces que lo que existe es materia (o energía), entendida esta como un tipo de existencia que ocupa una región espacio temporal, etc. La ciencia física se asociaría con un modo de darse de esa realidad material.

Por otro lado, una teoría como este materialismo postula una serie de entidades con ciertas características y afirma que son las únicas que existen. ¿Cómo saber si es verdad? Creo que, muy a la manera de las hipótesis científicas, con lo mejor que podemos quedarnos es con buenas razones para pensar que es verdad, con una plausibilidad creciente que no alcanza el 100%. No se prueba el fisicalismo. Esto es exactamente lo que sucede con teorías como el ‘Big-Bang’ o la teoría de la evolución: estamos tan persuadidos por la evidencia de que son verdaderas que simplemente asumimos que son hechos, pero es lógicamente posible que resulten ser falsas.

Las preguntas por el significado de “físico” y sobre la verdad del fisicalismo (o materialismo), parecen caminos prometedores para investigaciones futuras sobre el tema que he examinado en esta tesis, pero su éxito dependerá por completo de una valoración cuidadosa de asuntos que han venido quedando a lo largo de esta investigación, como *concebibilidad* y *posibilidad* (que incluye temas pertenecientes a la lógica modal y la

semántica de mundos posibles, e incluso relacionadas con teorías de la referencia y el significado), cerramiento del mundo físico, sobredeterminación causal; asimismo, la referencia de los términos de clase natural, la noción de explicación o el carácter explicativo de los enunciados de identidad, la designación rígida, entre otras. Espero que proyectos de investigación venideros puedan acometer tal labor.

Bibliografía

BLOCK, Ned (1980) 'Troubles with Functionalism', En: Readings in the Philosophy of Psychology, 1980.

BLOCK, Ned y STALNAKER, Robert (1999) 'Conceptual Analysis, Dualism, and the Explanatory Gap', The Philosophical Review, 108, pp. 1-46.

CHALMERS, David (1995) 'Facing up to the problem of consciousness', Journal of Consciousness Studies, 2 (3), pp. 200-219

CHALMERS, David (1996) The Conscious Mind. New York: Oxford University Press.

CHURCHLAND, Paul (1985) 'Reduction, Qualia, and the Direct Introspection of Brain states', Journal of Philosophy, 82, pp. 8-28.

DENNETT, Daniel (1991) Consciousness Explained. Boston: Little, Brown and Company.

DESCARTES, René (1977) Meditaciones metafísicas, Ediciones Alfaguara, Madrid.

HARMAN, Gilbert (1990) 'The Intrinsic Quality of Experience', En: J. Tomberlin (ed.), Philosophical Perspectives, Northridge, CAL: Ridgeview.

HORGAN, Terence (1984) 'Jackson on Physical Information and Qualia', The Philosophical Quarterly, 34, pp. 147-52.

JACKSON, Frank (1982) 'Epiphenomenal Qualia', The Philosophical Quarterly, 32, pp. 127-136.

JACKSON, Frank (1993) 'Armchair Metaphysics', En: J. Hawthorne y M. Michael (eds.) *Philosophy in Mind*, Amsterdam: Kluwer.

KRIPKE, Saul (1972) 'Identity and Necessity', En: Milton Munitz (Ed.) *Identity and Individuation*, New York: New York University Press, pp. 135-164

KRIPKE, Saul (1980) *Naming and Necessity*. Cambridge, MA: Harvard University Press.

LEVINE, Joseph (1983) 'Materialism and Qualia: The Explanatory Gap', *Pacific Philosophical Quarterly*, 64, pp. 354-361.

LEVINE, Joseph (1993) 'On Leaving out what it is like', En: M. Davies y M. Humpreys (eds.) *Consciousness Psychological and Philosophical Essays*, Cambridge, MA: Blackwell.

LEWIS, David (1988) 'What Experience Teaches', En: Lycan, W. (ed.) *Mind and Cognition: A Reader*. Cambridge, MIT: Blackwell, pp. 499-519.

LEWIS, David (1971) 'An Argument for the Identity Theory', *The Journal of Philosophy*, 63, pp. 17-25.

LEWIS, David (1980) 'Philosophical and Theoretical Identifications', *Australasian Journal of Philosophy*, 50, pp. 249-258

LOAR, Brian (1997) 'Phenomenal States: second version', En: N. Block, et al. (eds.) *The Nature of Consciousness: Philosophical Debates*, Cambridge, MA: MIT Press.

LYCAN, William (1995) 'A Limited Defense of Phenomenal Information', En: Metzinger, T. (ed.) *Conscious Experience*. Tucson: University of Arizona Press, pp. 243-258

MCGINN Collin (1991) *The Problem of Consciousness: Essays Towards a Resolution*. Oxford: Blackwell.

NAGEL, Thomas (1974) 'What is it like to be a bat', *The Philosophical Review*, 83 (4), pp. 435-450.

NEMIROV, Lawrence (1990) 'Physicalism and the Cognitive Role of Acquaintance', En: Lycan, W. (ed.) *Mind and Cognition: A Reader*. Cambridge, MIT: Blackwell, pp. 490-499.

PAPINEAU, David (1998) 'Mind the Gap', *Philosophical Perspectives*, 12, pp. 373-388.

PAPINEAU, David (2002) *Thinking about Consciousness*. Oxford: Oxford University Press.

PUTNAM, Hilary (1967) 'The Nature of Mental States', En: W. H. Capitan y D. Merrill (eds.) *Art, Mind and Religion*. Pittsburgh: Pittsburgh University Press. Pp. 37-48.

TYE, Michael (2000) *Consciousness, Color, and Content*. Cambridge, MA: MIT Press.

WIDER, Kathleen (1990) 'Overtones of Solipsism in Thomas Nagel's "What is it like to be a bat" and "The View from Nowhere"', *Philosophy and Phenomenological Research*, 3, pp. 481-499