

UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE FILOSOFÍA Y LETRAS

COLEGIO DE FILOSOFÍA

RUSSELL, REALISMO ESTRUCTURAL, MATERIA Y MENTE

T E S I S

QUE PARA OBTENER EL TÍTULO DE

LICENCIADO EN FILOSOFÍA

PRESENTA:

MOISÉS MACÍAS BUSTOS

TUTOR: MTRO. CRISTIAN ALEJANDRO GUTIÉRREZ RAMÍREZ

CIUDAD UNIVERSITARIA, D.F. JUNIO DE 2014



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

ÍNDICE

	Página
Dedicatoria	4
Agradecimientos	5
INTRODUCCIÓN	9
Russell, Materia y Mente	10
I. Trasfondo Histórico	10
II. La Posición Russelliana	14
III. Estructura del Trabajo	17
CAPÍTULO I: PROBLEMAS CON EL DUALISMO DE SUSTANCIAS	19
I. El argumento de la identidad personal a favor del dualismo.	20
II. El argumento de la prioridad metafísica de lo mental.	35
III. El argumento de la función.	38
IV. El argumento de la concebibilidad.	42
V. El argumento del conocimiento.	45
VI. Dualismo y Experiencia (<i>Qualia</i>)	50

CAPÍTULO II: PROBLEMAS PARA EL FISCALISMO REDUCCIONISTA	57
I. Argumentos en contra del fiscalismo	58
II. Algunas consideraciones sobre el problema mente-cuerpo	61
III. Algunos argumentos anti-fiscalistas explorados a fondo.	63
CAPÍTULO III: EL PROBLEMA DIFÍCIL DE LA MATERIA: LA TEORÍA DE BERTRAND RUSSELL	69
I. La Teoría Russelliana Causal de la Percepción	71
II. La Lógica Matemática de Russell y La Estructura del Mundo Físico	74
III. Espacio-Tiempo y Espacio Perspectival	83
IV. El Problema de Linderos	85
V. El Argumento Bidimensional en Contra del Fiscalismo	86
VI. En Síntesis	91
VII. Diagnósticos	93
CONCLUSIÓN	98
BIBLIOGRAFÍA	101

A Bertrand Russell

What is absolutely vital to me is the self-respect I get from work - when (as often) I have done something for which I feel remorse, work restores me to a belief that it is better I should exist than not exist. And another thing I greatly value is the kind of communion with past and future discoverers. I often have imaginary conversations with Leibniz, in which I tell him how fruitful his ideas have proved, and how much more beautiful the result is than he could have foreseen; and in moments of self-confidence, I imagine students hereafter having similar thoughts about me. There is a 'communion of philosophers' as well as a 'communion of saints', and it is largely that that keeps me from feeling lonely. – Excerpted from Ray Monk's *Bertrand Russell: The Spirit of Solitude*, p.

186.

AGRADECIMIENTOS

Este trabajo de investigación y mi formación filosófica en general se han visto no sólo beneficiados, sino posibilitados por la presencia de varias personas en mi vida.

Antes que nada quiero mencionar a mi familia, especialmente a mi madre, Libertad Bustos García, quien siempre ha tenido confianza en mí y cuyo apoyo ha sido una constante en mi vida. Es la mejor madre que alguien como yo podría tener. También mi hermano Leonardo Isaac Macías Bustos, quien es, para alguien como yo, el mejor hermano posible: inteligente, sensible, curioso y apasionado por sus temas. A María Fernanda Mora Alba, quien fue mi compañera de vida durante casi nueve años y con quien compartí algunos de los momentos más hermosos que he experimentado. Mi padrastro Luis Delgado Calderón quien con su constante buen humor y experiencia me ha dado muchas lecciones de vida. Mi tía Nidia Bustos García, cuya fuerte presencia siempre ha sido un pilar de apoyo para nuestra familia. También quiero agradecer a mis primos Nidia Gabriela Gutiérrez y Jesús Octavio Gutiérrez, así como a mi tío Jesús Gutiérrez Delgado, quien falleció hace poco dejando una honda tristeza entre quienes lo queríamos.

En segundo lugar quiero agradecer a mis amigos, quienes también han sido mis maestros e interlocutores. Principalmente quiero agradecer a mi amigo y supervisor Cristian Gutiérrez Ramírez. Sería imposible para mí en el espacio de unas oraciones el expresar la profunda admiración que siento por Cristian. Cristian es uno de esos casos más bien atípicos de personas que combinan una mente de primer nivel con un enorme corazón y un genuino compromiso por ayudar a los demás. A lo largo de los años me he beneficiado enormemente de discutir con él toda clase de temas filosóficos: desde filosofía de las

matemáticas, pasando por metafísica hasta meta-filosofía. Por un tiempo organizamos en mi casa un seminario informal de teoría de conjuntos avanzada el cual fue instrumental para que yo pudiera comenzar a interesarme y profundizar en temas de filosofía de la física en virtud de las complejas estructuras matemáticas que se abordan ahí. Temas que ahora se han vuelto mi principal interés filosófico. Por el lado puramente ético puedo decir que, al menos en mi caso, ser amigo de Cristian me hace intentar ser una mejor persona.

Mi gratitud con mis amigos no termina ahí. Todo el tiempo aprendo algo de ellos y me siento enormemente afortunado de tenerlos en mi vida. Genuinamente creo que son algunas de las personas más inteligentes y sensibles que caminan por el globo. Quiero agradecer a Jorge Alberto Treviño, Víctor Cantero Flores, Ricardo Mena Gallardo, Luis Estrada González, Javier García Salcedo, Teresa Verthein, Alejandro Vázquez del Mercado, Fernando Flores, Carlos Romero, Angélica Pena Martínez, Samuel Lomelí y Jesús Granados.

En tercer lugar, quiero agradecer a la comunidad de filósofos de Twitter, de la que hasta ahora he sido un miembro bastante activo. Muchas amistades ahí me han dado también valiosas lecciones: Allen Stairs, Stephen Mumford, Timothy Prisk, Guy Longworth y Anita Leirfall entre muchos otros.

Finalmente quiero agradecer a un par de profesores que, en mis interacciones con ellos, largas o cortas, me enseñaron mucha filosofía. Primero al Dr. Alejandro Tomasini Bassols, con quien tomé muchos excelentes cursos como estudiante de licenciatura. La enorme curiosidad del Dr. Tomasini nos llevó a los que fuimos sus alumnos a leer y discutir toda clase de textos, desde autores analíticos, hasta Kant y Plotino. En segundo lugar al Dr.

Agustín Rayo, cuya presencia en las aulas de la UNAM fue corta, pero cuyo rigor, facilidad expositiva y receptividad a discutir filosofía de alto nivel con los estudiantes dejaron una profunda impresión en mí. Puedo decir que su curso sobre Generalidad Absoluta ha sido, sin lugar a dudas, el mejor curso de filosofía que he tomado en mi vida.

Physics is mathematical, not because we know so much about the physical world, but because we know so little: it is only its mathematical properties that we can discover.

- Bertrand Russell, *An Outline of Philosophy*

INTRODUCCIÓN

El problema filosófico general que nos concierne en esta tesis es el de la naturaleza del mundo físico y la solución al problema mente-cuerpo, no en su aspecto intensional i.e. que los pensamientos puedan ser acerca del mundo¹, sino en su aspecto fenoménico i.e. el hecho de que hay algo que es tener experiencias².

En muchos trabajos el problema en torno a la naturaleza del mundo físico y nuestro acceso al mismo suele tratarse de forma separada al problema de la experiencia. Sin embargo, dado el enfoque en soluciones russellianas que guía este trabajo, voy a sostener que en la medida que nuestro conocimiento del mundo físico se deriva en primera instancia de nuestras percepciones y en segunda instancia de los modelos matemáticos que elaboramos para explicarlas, nada de lo que sabemos en física es incompatible con el hecho de que las propiedades fenoménicas (*qualia*) sean propiedades físicas.

Es importante entonces, de entrada, señalar cuales son los objetivos modestos de ese trabajo. No pretendo aquí decir cuál es la solución al problema mente-cuerpo sino meramente mostrar que existe una posición, en el espacio conceptual de teorías filosóficas sobre el tema, que muestra que es posible reconciliar a la física con la experiencia recuperando varios compromisos pre-teoréticos y sin postular un dualismo de sustancias.

Inclusive quiero señalar de entrada que existe ya una literatura creciente sobre posiciones russellianas en torno al problema mente-cuerpo y que si bien yo exploro un poco

¹ Jacob, Pierre, "Intentionality", *The Stanford Encyclopedia of Philosophy* (Fall 2010 Edition), Edward N. Zalta (ed.), URL = <<http://plato.stanford.edu/archives/fall2010/entries/intentionality/>>.

² Nagel, 1974.

el espacio conceptual que existe entre algunas de esas posiciones, no me comprometo aquí con ninguna sino solamente busco argumentar que el esquema general de la posición russelliana está en mucho mejor posición para abordar este problema de lo que lo están las posiciones tradicionales.

A continuación prepararé el terreno para abordar el problema repasando algunos de los enigmas y argumentos tradicionales en filosofía de la mente, metafísica y filosofía de la ciencia. Todo esto de forma que, más adelante, una vez establecido un trasfondo teórico, pueda hacer más inteligible el esquema de propuestas russellianas que tengo en mente.

-

Russell, Materia y Mente

I) Tránsito Histórico

Russell, a lo largo de varios escritos, esbozó una teoría sobre la física y sobre los eventos mentales que ofrece la posibilidad de superar algunas de las dificultades persistentes del problema mente-cuerpo. Aquí voy a tratar de delinear su posición general y partiendo de la misma voy a desarrollar una línea argumental que preserve el espíritu de su propuesta y muestre cuando menos que el fisicalismo russelliano está en una mejor posición que el ortodoxo para abordar los enigmas tradicionales acerca de la conexión entre mente y materia.

Uno de los aspectos más interesantes e importantes de la filosofía de Russell, una vez terminada la publicación de *Principia Mathematica*, fue el uso de la noción de

construcciones lógicas a partir de nociones lógicas y funciones proposicionales³, mediante la aplicación constante de su Teoría de las Descripciones, a las nociones más fundamentales de la física como él las comprendía, esto es: espacio, tiempo, causalidad y materia. Estas nociones se encuentran en el núcleo de su ulterior filosofía. Una vez equipado con un aparato formal poderoso, Russell sintió que podía proporcionar los fundamentos lógicos de la física sustituyendo inferencias sospechosas por construcciones lógicas.

Él enuncia su máxima metodológica de la siguiente forma:

Siempre que sea posible, sustituye construcciones de entidades conocidas a inferencias a entidades desconocidas.⁴

La idea detrás de este principio es aplicar una variante analítica de la navaja de Occam: es posible analizar entidades complejas, resolviéndolas en sus entidades componentes más simples y sus relaciones. Sustituyendo cualquier necesidad *de re* por necesidades lógicas⁵. Es evidente *prima facie* que la Teoría de las Descripciones de Russell⁶ tiene aplicaciones formales y lingüísticas. Russell vio, sin embargo, que el rango de aplicación de la teoría podía ampliarse para analizar otras clases de complejos, tanto en filosofía de la mente como en filosofía de la física. Su aplicación, por ejemplo, a la noción de ‘materia’ como la clase de todos aquellos *sense-data* y *sensibilia* en un punto desde toda

³ Recuérdese que las clases y relaciones en *Principia Mathematica* son a su vez analizables en términos de funciones proposicionales. Véase Linsky, 1999.

⁴ “Wherever possible, substitute constructions out of known entities for inferences to unknown entities.” *Logic & Knowledge*, p. 326.

⁵ Véase Landini, 2011.

⁶ Véase Russell, 1905.

perspectiva posible⁷. De esta forma Russell evita las inferencias, siempre dudosas, sustituyéndolas por construcciones lógicas que no agregan ni postulan nada sino que definen una cosa como función de otras más simples y de sus relaciones. Russell pensaba que este método, que había resultado fructífero en matemáticas y en filosofía de las matemáticas, podía aplicarse exitosamente a otras áreas del conocimiento científico. J.

Passmore nos da el ejemplo de la descomposición russelliana de un penique:

Así según Russell un penique por ejemplo consiste en el sense-datum elíptico que aparece en tu espacio privado y en el sense-datum redondo que aparece en el mío, así como en otros sense-data distintos de otros espacios privados. Estas diversas apariencias, dice, forman “una cosa”, en el sentido en que “se comportan con respecto a las leyes de la física de una forma en que en general no se comportarían series no pertenecientes a ella” – por lo menos esto es cierto de las “cosas” de las que se ocupa la física. En cuanto a las “cosas” del sentido común, Russell cree que se conciben con tan poca precisión que apenas es posible una explicación satisfactoria de su construcción.⁸

La razón por la cual Russell deseaba emplear el formalismo en el análisis de los conceptos de la física es que consideraba que tal análisis mostraría cómo es posible que la física sea una ciencia empírica. Esto es consistente con su deseo de vincular la ciencia a la experiencia. En la física es posible obtener predicciones que pueden ser confirmadas mediante los datos de la experiencia, de tal que estos últimos pueden comprenderse como estando en función de lo que indican nuestras mejores teorías. El procedimiento de Russell era el inverso, es decir, explicar cómo los objetos de los que habla la física están en función de los datos de la experiencia.

La manera en que Russell dio seguimiento a este proyecto tuvo modificaciones a lo largo de los años, así, durante el período de *Our Knowledge of the External World* (1914) Russell se suscribió a un fenomenalismo que tomaba en cuenta sólo los datos de la

⁷ *The Ultimate Constituents of Matter y The Relation of Sense-Data to Physics* (Russell, 1917).

⁸ Véase Passmore, 1981.

conciencia. Sin embargo, este fenomenalismo, ausente en *The Problems of Philosophy* (1911), está también ausente en *Sense-Data and Physics* (1913), *The Analysis of Matter* (1927) y *Human Knowledge* (1948). Libros que aquí serán nuestras fuentes principales.

En los enfoques alternativos Russell es más inclusivo con respecto a lo que necesitamos emplear para explicar cómo la física es una ciencia empírica. Así, en *Sense-Data and Physics* Russell se permite emplear *sensibilia*, esto es, *sense-data* que si bien se encuentran en alguna región espacio-temporal, no pertenecen a la perspectiva de ningún sujeto. Y en *The Analysis of Matter*, *Human Knowledge* y *The Problems of Philosophy* se vale de una teoría causal de la percepción.

El programa de construcciones lógicas de Russell influyó en la filosofía de los empiristas lógicos, notablemente Carnap⁹, pero se vio que había enormes dificultades si el proyecto tenía que partir exclusivamente de un fenomenalismo (los datos exclusivos de la conciencia). Por ello, enfoques como los de *The Analysis of Matter* han resultado ser más sólidos a la larga. En ese libro Russell emplea su método de construcciones lógicas para elucidar conceptos como: punto, materia, mente, instante y estructura. También avanza una tesis sobre el mundo físico, hoy en día muy discutida, conocida como *realismo estructural epistémico* que sostiene que del mundo físico sólo podemos conocer su estructura abstracta y que va a resultar fundamental a la discusión de esta tesis.

⁹ Véase el artículo de Christopher Pincock *Carnap, Russell and The External World* en *The Cambridge Companion to Carnap*, 2008.

II) La Posición Russelliana

En la literatura contemporánea en filosofía de la mente existen diversas objeciones a la posición conocida como fisicalismo, esto es, la teoría que sostiene que todo lo que existe en el universo es físico. Estas objeciones generan ciertas dificultades para teorías funcionalistas¹⁰ y teorías de identidad eliminativistas¹¹¹² como las de Smart o Place pero que, argumentaré, pueden ser superadas por posiciones Russellianas.

Respecto al problema mente-cuerpo Russell defendió en su filosofía tardía la tesis del monismo neutral i.e. la tesis que sostiene que la ontología fundamental del mundo no es mental, ni física sino que todo evento puede ser ya sea mental o físico dependiendo del tipo de estructuras en las que entre. Esta teoría está presentada por primera vez en *On Propositions: What they are and how they mean* (1919). Y se elabora de manera más robusta y sofisticada en *The Analysis of Mind* (1921), *The Analysis of Matter* (1927a) y *An*

¹⁰ Por funcionalismo entiéndase la teoría metafísica que sostiene que tener mentalidad es simplemente implementar una estructura causal-funcional a nivel físico (cerebros/autómatas) tal que, a ciertos estímulos, dada esa estructura, se dan, todo lo demás constante, ciertas respuestas. Para el funcionalista hay estados internos que, dada su estructura y ciertos estímulos, llevan *ceteris paribus* a ciertas respuestas. Existen variantes no-reduccionistas de funcionalismo que no pretenden decir que la experiencia es simplemente estructura causal-funcional. Véase Chalmers, 1996.

¹¹ Véase Smart, J. J. C., "The Mind/Brain Identity Theory", *The Stanford Encyclopedia of Philosophy* (Winter 2012 Edition), Edward N. Zalta (ed.), URL = <<http://plato.stanford.edu/archives/win2012/entries/mind-identity/>>.

¹² Las teorías de la identidad eliminativistas sostienen en esencia que tener cierto estado mental no es otra cosa que tener cierto estado cerebral i.e. los estados mentales son idénticos a estados cerebrales; pero tradicionalmente son los predicados mentales los que en estas teorías son parafraseados a predicados fisicalistas. En principio una teoría de identidad podría inclinarse al idealismo i.e. la tesis de que todo en el universo es mental. Para una defensa contemporánea del idealismo véase Foster, 1982.

Outline of Philosophy (1927b). Finalmente culminando en *Human Knowledge: Its Scope and Limits (1948)*.

Presentados en estos libros hay toda una serie de innovaciones que tienen en cuenta avances en la ciencia más avanzada, esto es, la física. Por ejemplo, Russell presenta una ontología del mundo en términos de eventos de un espacio-tiempo tetradimensional como el descrito por la teoría general de la relatividad i.e. una variedad diferencial de curvatura-variable con una métrica de tipo Minkowski, así como las leyes físicas para explicar la dinámica de los objetos construidos a partir de tales eventos.

A partir de ello desarrolla un novedoso enfoque al problema mente-cuerpo que va de la mano con su realismo estructural epistémico¹³. La idea central aquí es que del mundo físico sólo podemos inferir legítimamente su estructura abstracta. Pero si esto es así, entonces sólo conocemos el esqueleto causal del mundo y eso es compatible con que las cualidades que instancian esa estructura sean fenoménicas o proto-fenoménicas. Parafraseando a Michael Lockwood ‘la conciencia es de ellas [*qualia, sense-data*, etc], no con ellas’¹⁴.

El punto central de esta teoría, es que Russell piensa que lo que es conocido íntimamente es lo fenoménico, como opuesto a lo material. La mayoría de los teóricos de la identidad de la época afirmaban que los estados mentales son idénticos a estados cerebrales, y que los segundos nos son perfectamente cognoscibles a través de la fisiología, la observación empírica, etc. Russell, sin denostar la investigación empírica, estuvo siempre

¹³ La tesis del realismo estructural óptico, defendida entre otros por James Laydman (2007), es aún más fuerte. No sólo es que podamos del mundo conocer a lo más su estructura abstracta, sino que las entidades físicas son sólo estructuras.

¹⁴ Lockwood, 1989.

convencido que si algo desconocido a grandes rasgos es la materia. La mente por el contrario, al menos la propia, es aquello a lo que tenemos acceso privilegiado.

En la medida en que es legítimo el discurso sobre objetos materiales es que en tanto hipotéticos nos permiten una mejor reconstrucción de nuestra experiencia pese a estar más allá de toda experiencia posible. Por ello Russell afirma famosamente que lo que el fisiólogo ‘ve’, cuando ve otro cerebro es su propio cerebro¹⁵. Tomado literalmente esto es falso, pero como bien señala Lockwood (1989) Russell usa aquí ‘ver’ de una manera especial, posiblemente a propósito para irritar a otros filósofos. Lo que él quiere decir es que todos los eventos que percibe un individuo, ya sea que tengan vínculos con el mundo o sean acerca del mundo, en el sentido en que esos eventos son el eslabón final en una cadena causal, son eventos mentales que ‘son’ su cerebro, es decir, estos eventos son idénticos a alguna parte de su cerebro.

Como Russell era un monista neutral lo que esto querría decir es que estos eventos nos son intrínsecamente revelados como percepciones, sensaciones, etc; mientras que para un observador externo (extrínsecamente) serían la manifestación física de un cerebro. En un sentido nadie sabe exactamente cuales son las cualidades intrínsecas de los objetos físicos, con excepción de su propio cerebro, ya que estos están definidos o caracterizados, en la física y el lenguaje natural, por las propiedades físico-matemáticas que aparecen en la física o cualidades disposicionales. En vista de esto conocemos al menos las cualidades intrínsecas de al menos una porción del universo, nuestro cerebro aun cuando nuestro conocimiento del resto del mundo físico está limitado a su estructura.

¹⁵ Russell, 1959.

Esta sofisticada teoría no fue bien recibida en su momento aunque el espíritu de la misma fue defendido por, entre otros, Herbert Feigl y Grover Maxwell. Ha cobrado recientemente tremenda fuerza en debates de filosofía de la mente al grado de discutirse en ciertos círculos como una posible solución al problema mente-cuerpo¹⁶.

-

III) Estructura del Trabajo

El plan del trabajo es el siguiente. La tesis está dividida en tres capítulos. En el capítulo I exploro diversos argumentos en favor del dualismo de sustancias; desde argumentos que van de la identidad personal, pasando por la prioridad metafísica y la complejidad funcional de las personas hasta el problema de la experiencia. El objetivo de este capítulo es argumentar contra el dualismo, como una solución plausible al problema de la experiencia, tanto de forma indirecta como directa. La argumentación indirecta consiste en mostrar la redundancia del dualismo para todo tipo de problemas conectados con la metafísica de lo mental. Si el dualismo no es explicativo, ni útil en esos dominios entonces eso mostraría que su introducción para solventar el problema de la experiencia es completamente *ad hoc*. La argumentación directa pretende mostrar que el dualismo, en la mayoría de sus versiones, no nos ayuda tampoco a explicar el problema de la experiencia.

En el capítulo II expongo algunos de los argumentos más problemáticos, desde el punto de vista del problema de la experiencia, para el fisicalista. Señalo diferentes supuestos pre-teóricos a los cuales estamos comprometidos la mayoría de nosotros en el tema de la metafísica de lo mental y que considero toda teoría debe o acomodar o rechazar solamente con una justificación fuerte. Exploro diferentes maneras de reconstruir los

¹⁶ Chalmers, 1995.

argumentos de forma que se vean más claramente los diferentes problemas que tiene el fisicalismo tradicional al intentar lidiar con la experiencia. Finalmente discuto que estos argumentos muestran que hay un hueco explicativo insalvable desde el fisicalismo tradicional y que sería una buena estrategia repensar el fisicalismo desde una óptica russelliana.

En el capítulo III presento la teoría russelliana, primero basándome en las fuentes primarias y más adelante discutiendo fuentes secundarias así como algunas especulaciones personales. Paso de una discusión de su teoría causal de la percepción a una reconstrucción de los aspectos relevantes de su lógica-matemática; una reconstrucción de su posición respecto a los espacios de la física y la subjetividad; una discusión de un experimento mental queideo para establecer la coherencia de su posición y finalmente un argumento adicional, el argumento bidimensional de Chalmers, para fortalecer la posición dadas las discusiones previas. Empero, es necesario señalar que la adopción de una semántica bidimensional no es necesaria para adoptar una posición russelliana en la metafísica de la mente.

Finalmente hago algunos diagnósticos en torno a los supuestos pre-teoréticos, los argumentos anti-fisicalistas y argumento que la posición russelliana nos deja mejor parados respecto a los mismos que las posiciones tradicionales i.e. dualismo y fisicalismo ortodoxos.

Las traducciones a lo largo del texto son mías.

CAPÍTULO I

PROBLEMAS CON EL DUALISMO DE SUSTANCIAS

En ese capítulo se presentarán una serie de argumentos que se han avanzado a favor del dualismo de sustancias¹⁷ con el objetivo de mostrar la redundancia explicativa de ese tipo de dualismo respecto a temas como: función; prioridad metafísica y persistencia de las personas, así como diversos problemas que surgen al intentar usar tal tesis para dar cuenta de la experiencia. Se notarán diversas dificultades y se verá que, si bien el dualismo de sustancias no puede descartarse como lógicamente incoherente o como una real posibilidad para explicar los aspectos cualitativos de la experiencia, consideraciones del tipo navaja de Occam y de tipo ideológico¹⁸ nos inclinan a buscar otras alternativas.

Dentro de los argumentos a discutir habrá algunos clásicos (conocimiento, concebibilidad) y algunos que no se encuentran típicamente en las discusiones de filosofía de la mente sobre el dualismo de sustancias e.g. conexiones con identidad personal y prioridad metafísica.

Usaré el término “alma” para designar esa sustancia puramente mental en línea con la tradición.

¹⁷ El dualismo de sustancias en este contexto no es otra cosa que la tesis de que existen dos sustancias fundamentales i.e. básicas e independientes la una de la otra: materia y mente. Véase Wilson, 2012.

¹⁸ Por ideología entiendo en este contexto aquellos predicados, relaciones, cuantificadores, etc. que permiten, dentro de una teoría metafísica o científica, dar una explicación satisfactoria de la naturaleza del mundo. Por lo general se acepta que disminuir los compromisos ideológicos manteniendo el poder explicativo es una ganancia y podría ser evidencia de que esos predicados, relaciones y cuantificadores corresponden a la estructura última de la realidad (Sider, 2011). No pienso adentrarme en esa discusión.

I. El argumento de la identidad personal a favor del dualismo

Una línea argumental que podría parecer plausible dentro del espacio de soluciones del problema mente-cuerpo es aquella en la que, habiendo aceptado la existencia de sustancias mentales, las usamos para resolver una serie de acertijos conectados con el problema de la identidad personal. Si resolver este último problema pudiera motivar un dualismo de sustancias sería ventajoso usar esa sustancia mental para intentar también dar cuenta del problema de los *qualia* sin preocuparnos demasiado acerca de consideraciones ligadas con la navaja de Occam (ontología) o de tipo de ideológico.

Un típico ejemplo de los aspectos problemáticos de la noción de identidad personal es el siguiente:

Usualmente solemos pensar que el niño de la fotografía en nuestra habitación somos nosotros hace varios años. Igualmente, pensamos que en unos diez años estaremos más cerca de la vejez y otras cosas que presuponen la identidad personal, esto es, que una persona conserva su identidad a lo largo de la vida.

Al hablar de identidad personal la noción de identidad que usualmente se tiene en mente es la de identidad metafísica o matemática. Una manera de caracterizar esa relación es como aquella que un objeto tiene consigo mismo y con ningún otro. Una condición necesaria de la relación de identidad, más no suficiente, es que sea una relación de equivalencia sobre (plausiblemente) el dominio irrestricto i.e. el dominio más inclusivo de todos¹⁹. Una relación es de equivalencia sobre un dominio D si y sólo si es:

¹⁹ Véase Uzquiano y Rayo, 2007.

- i) Reflexiva: $\forall x (x = x)$
- ii) Transitiva: $\forall x \forall y \forall z ((x = y \ \& \ y = z) \rightarrow x = z)$
- iii) Simétrica: $\forall x \forall y (x = y \rightarrow y = x)$

Esta noción de identidad usualmente se maneja en lógica, lógicas de primer orden y de órdenes superiores así como teoría de conjuntos y evidentemente en el resto de las matemáticas. Un uso en la praxis del símbolo de identidad en estos lenguajes formales es el indicarnos cuando podemos sustituir un término por otro; metafísicamente hablando nos indica que un objeto del dominio, quizá bajo diferentes descripciones, tiene consigo mismo esta relación puesto que en la medida en que, en virtud de la ley de Leibniz se sigue la substitutividad de los idénticos, entonces si dos o más términos refieren al mismo objeto, pueden ser sustituidos en contextos extensionales: en síntesis, en contextos donde lo que está en juego son proposiciones y no actitudes proposicionales.

La Ley de Leibniz, también conocida como el principio de indiscernibilidad de idénticos, al menos como se maneja en filosofía contemporánea, se simboliza como sigue en segundo orden con letras griegas indicando subconjuntos del dominio de cuantificación (o en un metalenguaje de primer orden donde las Φ 's son fórmulas del lenguaje objeto).

Ley de Leibniz: $\forall x \forall y (x = y \rightarrow \forall \Phi (\Phi x \rightarrow \Phi y))$

Dado que esta ley no es (demasiado) controversial²⁰, pues sólo dice que algo comparte consigo mismo todas sus propiedades, de la misma podemos derivar algunas dificultades para quien acepta que hay identidad personal. Para empezar, dado que alguien

²⁰ Véase para más discusión Hawthorne, 2006 p. 7

cuando niño y cuando adulto tiene, en ambas fases, propiedades distintas, entonces la ley de Leibniz implica que el niño y el adulto son diferentes: si la identidad personal requiere que esos objetos sean estrictamente idénticos tendríamos que concluir que por lo tanto no hay identidad personal²¹.

Otro problema que enfrenta quien defiende la identidad personal lo presenta el siguiente argumento: El cuerpo humano está compuesto de átomos. Pero el cuerpo gana y pierde átomos, es un objeto complejo, de manera que es probable que cada determinado periodo de tiempo los átomos que conforman un cuerpo sean totalmente diferentes a los átomos que lo conformaron tiempo antes. Pero ¿cómo podríamos decir de un cuerpo, que en dos tiempos diferentes es materialmente distinto, que es el cuerpo de la misma persona, bajo el supuesto de que una persona es sólo su cuerpo? De entrada parece misterioso.

Pero un momento, la cosa no es tan sencilla. El defensor de la identidad personal tiene abiertas varias posibilidades (la siguiente lista no pretende ser exhaustiva):

- i) **Tridimensionalismo:** Un objeto²² persiste a través del tiempo ‘durando’, esto es, estando del todo presente en diferentes tiempos. Bajo este modelo el objeto es el mismo en cualesquiera dos tiempos diferentes en la medida en que tiene

²¹ Una de las mejores discusiones introductorias que conozco sobre este problema viene contenida en el artículo *Temporal Parts* de Ted Sider, incluido en Sider, Hawthorne and Zimmerman 2008 p. 241.

²² Qué sea un objeto en las diferentes posiciones metafísicas que surgen al intentar dar cuenta de la persistencia no está fijado meramente por la posición relevante: es decir, dependerá de otros factores de fondo como qué tipo de teoría de conjuntos; axiomatización de la mereología o metafísica del espacio-tiempo adoptemos. Lo que posiciones como el tridimensionalismo o tetradimensionalismo nos proporcionan es un núcleo de nociones que intuitivamente tendría que cumplir el objeto para persistir; pero decir categóricamente qué es un objeto tridimensional o tetradimensional requiere de otros elementos de fondo. Véase por ejemplo Sider 2003, Simons 1987.

sus propiedades en tiempos: es (niño en t) y (adulto en t'). Ese tipo de propiedades no son incompatibles por lo que se evita la contradicción. Pensemos, por ejemplo, en un camino es que quebradizo en una parte y recto en otra, claramente eso no es inconsistente. El tridimensionalista no requiere de nada por encima de las partes para explicar la persistencia. Hay que notar rápidamente que aceptar la posición no nos dará inmediatamente criterios de identidad a través del tiempo²³, sólo explicará cómo es que la identidad es posible. Lo primero requiere algún criterio ya sea 'identidad psicológica' o 'identidad física', ambos pueden dilucidarse en conexiones de acuerdo a alguna ley causal.

- ii) **Tetradimensionalismo:** Un objeto persiste a través del tiempo 'perdurando', esto es, estando presente en diferentes tiempos en la medida en que tiene partes temporales en esos tiempos. Bajo este modelo lo que difiere son las partes temporales del objeto: los objetos físicos de la vida diaria son gusanos espacio-temporales, extendidos en el tiempo, como en el espacio y tienen partes, que también son objetos físicos. Una analogía es la siguiente: yo tengo partes espaciales, como mi brazo, que ocupa una región espacial menor que la que ocupa todo mi cuerpo. Una parte temporal mía ocupa el mismo espacio que yo, pero lo hace por sólo un instante o un intervalo. Del objeto en cuestión podemos

²³ Es importante tener criterios de identidad, desde luego, en la medida en que creamos que, metafísicamente hablando, los objetos de hecho persisten a través del tiempo. Estos criterios son en esencia las condiciones que deben cumplirse para que el objeto persista en el tiempo. Posiciones como el tridimensionalismo o el tetradimensionalismo no dan por sí mismas la totalidad de esos criterios, pero si son coherentes muestran que la persistencia es posible.

decir entonces que tiene una parte temporal que es un niño y otra que es un adulto y éstas difieren, pero el objeto tiene la propiedad de estar compuesto por ambas partes temporales, por lo que no se cae en la contradicción. Las partes temporales tienen sus propiedades *simpliciter* por lo que no hay necesidad de hablar de propiedades en tiempos; al menos en conexión con esto. El tetradimensionalista no requiere nada por encima de las partes para explicar la persistencia. Hay que notar rápidamente que aceptar la posición no nos dará inmediatamente criterios de identidad a través del tiempo, sólo explicará cómo es que la identidad es posible. Lo primero requiere algún criterio ya sea ‘identidad psicológica’ o ‘identidad física’, ambos pueden dilucidarse en conexiones de acuerdo a alguna ley causal.

- iii) **Teoría del Alma (Una versión)²⁴**: Esta posición es similar al tridimensionalismo. El objeto persiste a través del tiempo estando del todo presente en diferentes tiempos. Sin embargo en este modelo lo que cambia es el cuerpo de la persona, puesto que la persona es en realidad idéntica a su alma y el cuerpo sólo es un objeto físico que tiene una relación de cercana intimidad a su alma en la medida en que hay una interacción físico-psíquica.

²⁴ Las variaciones en detalle de la teoría del alma (sustancial) que tengo en mente se verán más abajo cuando comience a desglosar las diferentes maneras en que podría echarse a andar esa posición. Lo que estoy dando aquí es un esquema de algunas de las condiciones que tendría que cumplir esta variante de la teoría del alma si ha de dar cuenta de la persistencia.

iv) **Teoría Psicológica de la Identidad Personal:** Algo que podría señalarse es que alguien podría considerar que lo importante en la identidad personal no es, irónicamente, la identidad en su sentido matemático o metafísico, sino más bien la continuidad psicológica o algún otro tipo de continuidad. Bajo este modelo las razones que yo tengo para considerar al niño que fui o al anciano que seré como importantes es simplemente que hay, digamos, una continuidad psicológica entre nosotros de gran estrechez. Esta continuidad, además, no se presenta en ningún otro caso. Por ello es que no puedo evitar pensar al niño y al anciano como idénticos a mí. En realidad la persistencia, bajo esta postura, es una ilusión. Que una persona persista es simplemente que existe cierta similitud entre diferentes psicologías de diferentes cuerpos en diferentes tiempos. Si estos cuerpos están conectados por una ley causal aún mejor, pero en cualquier caso lo que da la apariencia de continuidad es la cercanía psicológica a través del tiempo.

El bosquejo de posiciones de arriba, si bien no exhaustivo, permite al menos atisbar que es posible una solución al problema de la persistencia a través del tiempo. Sin embargo entrar en discusión respecto al tetradimensionalismo y tridimensionalismo, o explorar en detalle la teoría psicológica, nos alejaría de nuestro tema por lo que no ahondaré demasiado más allá de hacer un breve diagnóstico.

Podemos decir respecto al tridimensionalismo que tiene varias dificultades; entre éstas:

- i) No es claro cómo pueda caracterizarse. Dado que obviamente el tridimensionalista deseará aceptar que los objetos ganan y pierden partes pero bajo alguna formulación del tridimensionalismo como: ‘un objeto es tridimensional si y sólo si **necesariamente** (está del todo presente en todo tiempo en el que existe)’ entonces cualquier objeto que pierda o gane partes no es tridimensional²⁵.
- ii) *Prima facie* el tridimensionalismo nos compromete al abandono de propiedades intrínsecas, esto es, propiedades que un objeto tiene independientemente de otros objetos, dado el argumento de intrínsecos temporales de David Lewis (1986). El argumento en breve es que plausiblemente queremos mantener que algunas propiedades son intrínsecas a los objetos e.g. su forma. Que una propiedad sea intrínseca al objeto quiere decir que es posible que el objeto pudiera tener la propiedad aun si fuera el único objeto en el mundo. Sin embargo bajo el tridimensionalismo las propiedades siempre están relativizadas a tiempos. Un objeto no es esférico *simpliciter* sino (esférico en t), por decir. El argumento entonces iría: El tridimensionalismo implica que los objetos materiales no tienen propiedades intrínsecas; pero las tienen, ergo el tridimensionalismo es falso.
- iii) Si el argumento de la vaguedad de Ted Sider (2001) es exitoso entonces también muestra que el tridimensionalismo nos compromete con la tesis de que

²⁵ Para una extensa y sutil discusión de lo problemático que es caracterizar el tridimensionalismo véase el artículo de T. Sider *Fourdimensionalism; Philosophical Review* 1997.

hay vaguedad en el mundo (ontológica). La idea es sencilla, en mereología clásica²⁶ es un axioma que para cualesquiera dos objetos diferentes A y B, existe un único objeto C, la fusión de A y B, que tiene como partes propias únicamente a A, B y cualquier parte propia de A o B si éstos no son simples i.e. objetos sin partes propias. Sin embargo hay quienes consideran que la mereología clásica es demasiado fuerte para caracterizar la relación de todo-parte de los objetos materiales. Ellos querrían que cualesquiera objetos tuvieran una fusión sólo en condiciones especiales: es decir tienen en mente un esquema como: A y B se fusionan en C si y solamente si _____. Donde la línea en blanco captura alguna condición metafísicamente relevante. Sider (2004) argumenta que las típicas condiciones metafísicas plausibles para fusiones trans-temporales i.e. para objetos que existen en diferentes tiempos, son tales que, para algún tiempo t es indeterminado metafísicamente cuantos objetos materiales hay en el mundo²⁷. Pero eso no es posible, mantiene. Por lo tanto, o la composición mereológica trans-temporal es clásica i.e. cualesquiera objetos en diferentes tiempos tienen una fusión y son las partes temporales de esa fusión o no hay objetos complejos. Pero hay objetos complejos, por lo tanto cualesquiera objetos en diferentes

²⁶ La mereología es la teoría formal del todo y la parte. Véase Simons 1987.

²⁷ Por ejemplo, supongamos que dos objetos A y B trans-temporales tienen una fusión C si y solamente si ambos están vivos en esos tiempos. Pero plausiblemente hay instantes entre la vida y la muerte en donde es indeterminado si el cuerpo está vivo o muerto en la medida en que la aplicación de esos criterios es vaga. Se seguiría que para uno de esos tiempos es indeterminado si existe la fusión C. Y esa indeterminación no podría deberse a la vaguedad del lenguaje, porque cuantos objetos concretos hay en el mundo puede expresarse en un lenguaje de primer orden en donde ni los cuantificadores, ni las constantes lógicas ni el predicado C de concreto son susceptibles de ser vagos. Por lo que sería indeterminado cuantos objetos contiene el mundo en algún tiempo. Para más detalles véase Sider 2004.

tiempos tienen una fusión y son las partes temporales de esa fusión lo cual, piensa Sider, es suficiente para garantizar la verdad del tetradimensionalismo.

Por estas y otras consideraciones es razonable pensar que el tridimensionalismo es una posición implausible.

Una ventaja que tiene la teoría del alma respecto al tridimensionalismo es que permite señalar que, sin importar si el cuerpo gana o pierde partes, el alma no puede ganar ni perder partes en la medida en que, preferiblemente, el alma debe postularse como un objeto mereologicamente simple i.e. sin partes propias. Si el alma fuese compleja entonces, en la medida en que tiene partes, es concebible que el alma se descompusiera en sus partes, dejando de existir o impidiendo generar la simpleza deseada para la persistencia. Pues, presuntamente, lo que el defensor de la identidad personal quiere es decir que lo que típicamente llamamos persona, esto es un cuerpo humano que tiene condiciones claras de individuación, es el mismo ente desde el nacimiento hasta la muerte. Pero defender esto sólo es posible en la medida en que el alma es simple: si el alma es compleja nada garantiza que no pudiera descomponerse en sus partes cuando el correspondiente cuerpo está a la mitad de su vida, de forma que el intervalo durante el cual existe el cuerpo que típicamente llamamos persona excedería al intervalo durante el cual existiría la persona si la persona fuese idéntica a esta alma compuesta.

La razón por la cual el alma, en tanto se asume simple, podría explicar la identidad personal es que el alma no cambiaría en la medida en que no adquiriera o pierda partes; el alma es indestructible si aceptamos que la destrucción consiste en separar las partes, etc. Si la persona es idéntica a su alma parece que no puede ser diferente en t y en t' , más bien es

el cuerpo que está ligado al alma lo que es diferente. Por lo que el alma de Sócrates, digamos, no cambia aun cuando su cuerpo sí, pero él es idéntico a su alma, por lo tanto la ley de Leibniz no nos fuerza a concluir de Sócrates en t y en t' que no es idéntico a sí mismo.

Empero, hay al menos tres líneas críticas que se pueden delinear contra esta posición:

La primera es que, si las almas son simples, no es claro como dos almas podrían diferir más allá de numéricamente (es decir, diferir en cualidades intrínsecas i.e. que posean independientemente de todo otro objeto). Parece que una buena cantidad de propiedades sólo pueden tenerse cuando hay complejidad mereológica, esto es, cuando los objetos tienen partes. Por ejemplo, un auto tiene la propiedad de ser azul cuando todas o casi todas sus partes externas son azules. Pero el alma no puede tener ese tipo de propiedades al ser espacialmente simple. Por otro lado, no hay buenas razones para pensar que el alma puede tener propiedades mentales, aunque esta línea de argumentación la desarrollaré más abajo. Parece que quizá el alma podría poseer propiedades como las que poseen los objetos abstractos, como los números, los cuales *prima facie* no tienen partes espaciales. Sin embargo no es claro qué tipo de propiedad podría poseer el alma como un objeto abstracto, además de que, si las hubiere, no parecen ser el tipo de propiedades que nos ayudarían a entender la identidad personal (y mucho menos otras propiedades, como propiedades morales, que interesan al defensor del alma). Es relevante aquí notar que si las diversas almas no tienen propiedades intrínsecas, entonces es un poco raro decir que es el alma la que es idéntica a la persona. Puesto que claramente lo que nos hace distinguir a unas

personas de otras son sus conductas, apariencias, pensamientos, etc. Es claro que el alma simple no puede tener ninguna de esas cosas, las cuales son buenos candidatos para ser propiedades intrínsecas.

En segundo lugar, supongamos que deseamos mantener al alma como un objeto concreto, de forma que pueda conectársele fácilmente con el cuerpo.

Si el alma ha de ser simple y concreta, sólo se me ocurre asumir que el alma ocupa un sólo punto espacial, el cuál por definición no tiene partes. Quizá ocupa siempre un punto en el interior de un cuerpo físico. Esto presenta varios problemas. Por un lado presupone una filosofía del espacio ni mucho menos evidente: el sustantivismo, i.e. la tesis de que el espacio y/o el tiempo existen independientemente de los objetos que los ocupan o inclusive que éstos no son sino modos del espacio-tiempo (monismo sustantival), de forma que pueda decirse que existe una relación de ocupación entre el alma y el espacio (compuesto por puntos). Los relacionistas i.e. aquellos que sostienen que el espacio-tiempo no es fundamental sino reductible o fundamentado en las relaciones espacio-temporales de objetos materiales, pueden también construir un espacio con puntos a partir de nociones de lógica matemática y mereología, pero, hasta donde sé, los puntos que podrían proveer son sólo relativamente simples (en torno a algún dominio de objetos materiales) pero nada garantiza su simplicidad absoluta (véase, por ejemplo, Russell, 1927a). Es por ello que el sustantivismo es la única teoría que puede garantizar la simpleza en los puntos.

Por otro lado, sigue teniendo el problema de que no es claro cómo algo así podría tener propiedades que la distingan de otras almas, más allá de propiedades numéricas (i.e. que difiere de éstas). Supongamos que estas propiedades son relacionales, i.e. que está

siempre dentro de cierto cuerpo, a cierta distancia de la superficie de la Tierra, etc. Esto puede distinguir un alma de otras, sin embargo esto sólo empeora las cosas. Es claro que si el mundo cambia en lo más mínimo de un instante a otro, entonces las propiedades relacionales del alma cambiarán. Por ejemplo, si yo me muevo cinco metros de t a t' , entonces mi alma, la cual es espacialmente simple, dado que sus relaciones (espaciales) con los demás objetos han cambiado de t a t' , no tiene en t' las mismas propiedades que en t (las propiedades relacionales también son propiedades), y por lo tanto, el alma simple sería diferente en t y en t' (por Ley de Leibniz). ¡Vaya ayuda! Ahora el alma que iba a hacer el trabajo de explicar identidad personal nos genera un problema análogo de identidad del alma a través del tiempo. Parece que el alma no cambiaría propiedades sólo en mundos donde el sustantivismo sobre el espacio-tiempo valga y donde no hay cambio por periodos extendidos de tiempo (por lo que no habría cambio en el mundo durante ciertos intervalos). Pero en esos casos el cuerpo tampoco cambiaría, por lo que inclusive en esos mundos el alma no sería de mucha ayuda para explicar identidad personal, puesto que podríamos mejor valernos del cuerpo, cuya existencia es menos dudosa.²⁸

Una tercera objeción es que es concebible que inclusive un objeto perfectamente simple pudiera ser destruido, si entendemos destrucción no como descomposición en partes sino como desaparición total en algún tiempo. Pensemos en un sujeto, llamémosle Juan. Si ésta teoría del alma es correcta, entonces podría ser que el alma de Juan, sin que el pobre lo

²⁸ Aunque, a fin de cuentas, pienso que la teoría de identidad personal más plausible debe estar en la vecindad de la que Lewis desarrolla en *Survival & Identity*, una teoría tetradimensionalista de la identidad personal o en su defecto una teoría como la de Robert Nozick *The Closest Continuer Theory* (Nozick, 1981). Pues aun cuando esta última no preserva identidad metafísica si combina armoniosamente la teoría psicológica con la, aquí no discutida, teoría material de la identidad personal.

sepa, fuera destruida en algún punto. Pero entonces resultaría que el cuerpo de Juan sigue haciendo todas sus actividades, se casa, tiene hijos, avanza en su profesión medianamente pese a ser sumamente competente, (pues vive en un país corrupto), etc. Todo esto, pese a que su alma no existió durante la mayor parte de su vida (o de la vida del cuerpo materialmente pero no espiritualmente animado de Juan). Esto es sumamente implausible.

Definitivamente ligado a esto, podría ser, bajo esta teoría, que cada uno de los miembros de la humanidad desde hace mucho tiempo, digamos desde la revolución sexual, hubiera perdido su alma y sin embargo continúan con sus actividades, deseando vivir, creando, pensando, destruyendo, etc.²⁹ ¿No es una teoría con esas consecuencias demasiado increíble?³⁰ Por otro lado, es completamente inútil para explicar cómo el niño en la fotografía es idéntico a mí en la medida en que, de acuerdo a esta teoría, el cuerpo definitivamente va a atravesar por cambios aun cuando no tenga alma, entonces si yo perdí mi alma por alguna imperdonable fechoría entre la infancia y la madurez, no puedo ser idéntico a ese niño. Esto choca demasiado con nuestras intuiciones, que son las que parcialmente queríamos rescatar con una teoría de la identidad personal. Todas estas

²⁹ Una posibilidad sugerida por Cristian Gutiérrez (comunicación personal) es que sea metafísicamente imposible, si bien lógicamente posible, que las almas no puedan ser separadas de sus correspondientes cuerpos hasta la muerte biológica de estos. Considérese por analogía la posibilidad de los simples extendidos que se discute en metafísica contemporánea i.e. objetos que no son metafísicamente indivisibles aun si su división es lógicamente posible. Si bien concedo que el espacio conceptual alberga esta posibilidad, me parece un movimiento demasiado *ad hoc* en el contexto de esta discusión. Aunque sin duda podría ser interesante considerar motivaciones teológico-filosóficas para una posición así.

³⁰ Es completamente obvio para cualquiera que tenga la más mínima formación en lógica que demostrar categóricamente la falsedad de una proposición requiere, bajo supuestos clásicos, derivar de la misma una contradicción. Pero desde luego la mayoría de argumentos filosóficos no pretenden algo tan fuerte respecto a posiciones rivales, meramente mostrar consecuencias indeseables, relativo a otras proposiciones que aceptamos i.e. mostrar que las posiciones rivales tienen ciertos costos teóricos.

consecuencias de la teoría son sumamente implausibles, pero concebibles dados los escenarios que hemos estado planteando. Luego, si la concebibilidad es evidencia de posibilidad lógica (y hay buenas razones para creer que lo es), parece que la teoría del alma, aun cuando ésta se asume simple, no nos garantiza la persistencia a través del tiempo.

Sin embargo, tanto el tetradimensionalismo como la teoría psicológica de la identidad personal ofrecen más ventajas. El tetradimensionalismo ofrece una lista larga de ventajas (Sider, 2004); empero aquí no es el lugar para entrar en ellas.³¹

Diremos al menos que bajo una plausible concepción tetradimensional una persona es un gusano espacio-temporal, un objeto que tiene partes espacio-temporales, y es idéntico a sí mismo en la medida en que es la suma de determinadas partes *sub specie aeternitatis*. El objeto cambia en la medida en que tiene diferentes partes temporales en diferentes tiempos, cada una de las cuales tiene propiedades *simpliciter* y donde una parte temporal del objeto es una parte del objeto que empalma³² todo lo que es parte del objeto en t y sólo en t . Así, podemos decir de alguna persona P , concebida típicamente como un mismo cuerpo físico bajo ciertas leyes causales como las de la física/biología, que es un objeto extendido en la dimensión espacial y temporal, desde su nacimiento en t , hasta su muerte, en t' y ese objeto tiene consigo mismo una relación que satisface las propiedades necesarias de la identidad i.e. una relación de equivalencia. Como ya se dijo, el tetradimensionalismo no postula nada por encima de las partes del objeto.

³¹ Russell es un tetradimensionalista respecto a los objetos materiales. Véase por ejemplo su 1924, 1927^a y 1927b.

³² X empalma a Y , en mereología clásica, si y solamente si existe un Z , tal que Z es parte de X y Z es parte de Y . (Simons, 1987)

Respecto a la teoría psicológica, es superior a la teoría del alma en la medida en que, como el tetradimensionalismo, no postula entidades adicionales y también en tanto permite rescatar ‘lo que importa’³³, respecto a nuestros juicios de identidad a través del tiempo.³⁴ Si bien bajo esta concepción no es necesario que el objeto satisfaga las propiedades de la identidad metafísica o matemática, sólo que preserve continuidad psicológica relevante. Aquí quizá es relevante notar que algunas versiones de tetradimensionalismo o tridimensionalismo pueden operar conjuntamente con una teoría de la contraparte temporal³⁵ en lugar de una teoría de identidad trans-temporal. En esos casos, es claro que la teoría psicológica es claramente compatible con esas dos posiciones.

No parece entonces que haya razones para suponer que el dualismo de sustancias nos ayude especialmente en el caso de la identidad personal. Parece que hay varias posturas preferibles y conceptualmente explicativas.

II. El argumento de la prioridad metafísica³⁶ de lo mental

³³ Parfit (1984) considera a la luz de diversos experimentos mentales que “lo que importa” en la supervivencia no es la identidad, sino cierto tipo de continuidad psicológica.

³⁴ Véase sin embargo Williams 1970, para otros problemas que enfrentaría la teoría psicológica.

³⁵ La teoría de la contraparte es una teoría desarrollada por David Lewis (1968) que pretende dar cuenta de la semántica de proposiciones que hacen alusión a mismidad de individuos en contextos modales sin introducir una modalidad primitiva. La relación de contraparte ahí postulada recupera, en ciertos contextos, hechos importantes acerca de nuestro discurso sobre individuos en diferentes mundos posibles pese a ser más débil que la relación de identidad. Haciendo uso de la teoría de contraparte uno podría abandonar la noción de que existe la identidad (metafísico/matemática) a través de mundos posibles. Sider (2004) y otros han extendido la teoría de la contraparte para uso en contextos donde la identidad trans-temporal es relevante.

³⁶ Entiéndase en el sentido de fundamentación (*grounding*).

Este argumento es ilustrativo de una serie de nociones y problemas relevantes³⁷. Por ello lo detallaré en breve y daré un par de razones por las cuales considero que no funciona.

Locke, en su *Essay Concerning Human Understanding* dice lo siguiente:

Si, entonces, hubiese algo eterno, veamos qué tipo de ser debiera ser. Respecto a ello es muy obvio a la razón, que debe ser necesariamente un ser pensante. Pues es imposible pensar que la materia no-pensante pudiera producir un ser inteligente pensante como si la nada, de sí misma, produjese la materia...³⁸

La idea es la siguiente: de la nada no puede originarse nada, pero suponiendo que hubiese materia, entonces, ya sea que ésta se mueve o no. Si se mueve, tuvo que haber sido puesta en movimiento o haberse estado moviendo por la eternidad, consistente con la física newtoniana³⁹. Pero concediendo que ha estado moviéndose desde la eternidad, ningún movimiento de la materia podría producir el pensamiento. Ergo, el pensamiento, dado que existe, antecede (metafísicamente) a la materia. En el argumento de Locke, es el pensamiento de Dios.

³⁷ Ésta discusión está básicamente inspirada en aquella que desarrolla Daniel Dennett en su *Darwin's Dangerous Idea* (1995, p. 26) en conexión a estas reflexiones de Locke. Sigo una línea similar pero enfatizo las conexiones relevantes con mi tema.

³⁸ "If, then, there must be something eternal, let us see what sort of Being it must be. And to that it is very obvious to Reason, that it must necessarily be a cogitative Being. For it is as impossible to conceive that ever bare incogitative Matter should produce a thinking intelligent Being, as that nothing should of itself produce Matter..." de *Darwin's Dangerous Idea*, p. 26.

³⁹ Los objetos en la física de Newton mantienen su estado inercial hasta que se les aplique una fuerza i.e. se mantienen en movimiento o en reposo. La estructura del espacio Newtoniano es tal que permite la existencia de posiciones absolutas y por ende permite distinguir: reposo absoluto, movimiento rectilíneo uniforme y aceleración. Véase Sklar 1977.

Es importante enfatizar que el tipo de mentalidad en juego en este argumento es la funcional⁴⁰. En tanto lo único requerido del ‘pensamiento’ para fines de lo discutido es su capacidad de ‘pensar’, ‘interactuar con lo material’, ‘discernir’, etc. Todas características de lo mental que tienen una dimensión puramente funcional.

Lo importante aquí no es la cuestión teológica, sino el hecho de que si efectivamente el argumento, un argumento del designio, estableciera la imposibilidad de generar mentalidad de lo material entonces mente y materia difieren, en tanto que tienen diferentes propiedades. El argumento presupone que lo material genera lo material trivialmente (se genera a sí mismo): es una propiedad de lo material el ser material, pero si lo material no genera lo mental, entonces lo mental tiene la propiedad de no ser generado por lo material. Propiedad que lo material sí tiene. De nuevo, por Ley de Leibniz podemos concluir su diferencia, si concedemos el resto de las premisas.

A éste argumento puede responderse que la teoría darwiniana efectivamente muestra, como argüiré abajo, que lo mental-funcional⁴¹ puede generarse a partir de lo material. Si bien habría de entrada que notar que en este esquema, lo mental-funcional no es algo que tenga todo objeto físico, sino sólo cierto tipo de objetos de los cuales podríamos decir, desde el punto de vista de la tercera persona, que tienen creencias, recuerdos, pensamientos, etc.

⁴⁰ Por aspecto funcional entiendo aquellas características de algún X que, en cierto complejo estructural, juegan un rol determinado. Por ejemplo, en mi mesa las patas de madera juegan el rol de soportar el peso del disco, pero las patas bien podrían ser de metal. Similarmente en estructuras complejas, tanto físicas, como computacionales o mentales, ciertos elementos juegan un determinado rol funcional. Es importante enfatizar que el que algo tenga propiedades funcionales no implica la tesis del funcionalismo.

⁴¹ Por “mental-funcional” entenderé aquellas características funcionales comúnmente asociadas con la posesión de una mente e.g. la capacidad de tener creencias, recuerdos, pensamientos, meta-cognición, etc.

Bajo el esquema darwiniano, recordemos en breve, tenemos evolución y selección natural. La evolución es el cambio que se da en las especies y el algoritmo para el cambio es el de selección natural. En esencia, aquellos seres mejor adaptados a su entorno tienden a sobrevivir y dejar descendencia.⁴² La adaptación para un organismo x consiste en que, dado el medio, x tiene las características que en competencia, estadísticamente, tenderán a favorecerle reproductivamente. Así, si gradualmente algún organismo desarrolla, por alguna mutación aleatoria, alguna característica o proto-característica que pueda caracterizarse más o menos como: estar alerta, responder a una gama más amplia de estímulos, etc⁴³. Esas características le conceden una ventaja (reproductiva) estadística en su medio, entonces sus descendientes a lo largo de mucho tiempo probablemente generarán capacidades similares.

Ese tipo de capacidades, observables desde un punto de vista de la tercera persona, son algunas de las que intuitivamente llamamos mentales, en el sentido de mental-funcional. Vemos entonces que al menos lo mental-funcional puede derivarse de lo material si se acepta la teoría darwinista. Sobre lo mental-experiencial⁴⁴, i.e. *qualia*, a menos que se presuponga que es reductible a lo mental-funcional entonces no se puede adelantar un diagnóstico; pero veremos en el capítulo II que bajo una concepción ortodoxa de la materia es difícil ver cómo derivar los *qualia*.

⁴² Véase Dawkins, 1976.

⁴³ Los así llamados efectos fenotípicos de los organismos y que son funciones de elementos genéticos y ambientales, están en última instancia fundamentados en la química y biología molecular y los efectos que elementos en esos dominios nómicos pueden tener en el medio. Véase Dawkins, 1982.

⁴⁴ Los aspectos de primera persona que solemos llamar también mentales, específicamente el tener experiencias. El que, tomando prestada la frase de Nagel (1974), “exista algo que sea ser” una persona u objeto de experiencia.

Sin embargo, en el contexto de esta discusión, no parece que haya fuertes razones para pensar que se requiere de alguna mentalidad metafísicamente anterior que genere la materia y luego la mente humana, en la medida en que lo que el argumento exigía de esa mentalidad eran características funcionales como podrían ser la inteligencia o cognición. Tentativamente esto bastará para hacerlo a un lado.

III. El argumento de la función⁴⁵

El defensor del alma todavía puede, en este punto, apelar al poder del alma para explicar la vida mental de las personas. Respecto a si el alma puede ayudar a explicar los aspectos mental-funcionales (de organización mental, memoria, capacidad de pensar, tener creencias, disposiciones, etc.), los prospectos creo, son pobres. Sin embargo consideraré brevemente un argumento intuitivo en favor de que hace falta más que lo material para recuperar esas capacidades.

El argumento que esbozaré a continuación se presenta de manera general en las clases que Shelly Kagan, profesor de la universidad de Yale, impartió hace unos años bajo el título '*Death*' y son accesibles desde el portal de dicha universidad.⁴⁶ Durante las clases 2, 3, 4 y 5 Kagan aborda el debate entre dualistas y fisicalistas⁴⁷ con una serie de argumentos que van desde lo más endeble hasta lo más robusto. El argumento de la

⁴⁵ Éste argumento lo discuto en prácticamente las mismas palabras en un post que apareció, bajo mi autoría, en el blog de difusión de la filosofía que inicié junto con mi colega Carlos Romero: La Pluralidad de los Mundos www.pluralidad.tk

⁴⁶ <http://oyc.yale.edu/philosophy/phil-176>

⁴⁷ Me disculpo si en el cuerpo del texto ocupó tanto el término "materialismo", como el término "fisicalismo" en la medida en que los uso como sinónimos: a saber, la tesis de que todos los objetos concretos son físicos.

función puede considerarse, entre esos argumentos, como teniendo una posición intermedia.

La idea es la siguiente: Los seres humanos somos capaces de llevar a cabo toda clase de ‘funciones’: podemos enamorarnos, cantar, reflexionar, componer música, tener afectos, disfrutar de un buen libro, etc. Pero ¿cómo podría un objeto meramente físico llevar a cabo todo ese tipo de funciones sofisticadas? Por lo general suponemos a la mente o lo mental como algo mucho más sofisticado que lo meramente material, pero si aceptamos la identidad entre lo mental y lo material entonces tendríamos que decir que, en alguna situación dada, tanto lo mental como material serán igualmente sofisticados en tanto son lo mismo.

Pero pensemos en casos de objetos físicos típicos, por ejemplo una piedra. ¿De qué tipo de funciones es capaz una piedra? No de demasiadas, la piedra puede moverse de acuerdo a las leyes de la física. Si está en un estado de movimiento dado permanecerá en ese estado hasta que se le aplique una fuerza. Consideremos los átomos, los cuales componen todo objeto físico conocido excepto partículas elementales (como los quarks o electrones); partículas que a su vez los componen. Si asumimos la física lo más que podemos decir de los átomos es que se moverán acorde a las leyes dinámicas fundamentales y en sus interacciones con otros átomos formarán compuestos químicos. ¿Cómo pueden objetos físicos así componer a un ser capaz de reflexionar, amar, descubrir, hablar, componer música, etc.?

Es crucial sin embargo notar todo lo que hemos dejado fuera de estas consideraciones preliminares. Para empezar quizá los átomos sólo puedan llevar a cabo

funciones muy básicas, sin embargo los elementos que componen tienen un nivel de complejidad muy superior. Varias moléculas de agua pueden congelarse, fluir, evaporarse, flotar, etc. Análogamente, en el caso de la piedra parece que escogimos un objeto particularmente simple para ejemplificar.

Kagan pone como ejemplo de un objeto físico más complejo un teléfono celular o un radio. Ciertamente un teléfono celular es un objeto físico, pero para nada un objeto físico como la piedra, es mucho más complejo y es capaz de toda clase de funciones. Kagan llama a las funciones que puede llevar a cabo un objeto físico P-funciones. ¿Cuál es la incoherencia en suponer que el cuerpo humano es capaz de P-funciones mucho más complejas que la piedra, el teléfono celular, el radio, etc.? Parece que no tenemos razones *a priori* para decir que un objeto físico no podría llevar a cabo funciones cada vez más sofisticadas.

Siguiendo la línea argumental de arriba podemos considerar otro ejemplo adicional. Tengo en mente a Asimo, un robot humanoide creado por Honda. Cualquiera puede ingresar a *Youtube* y observar a Asimo caminando, corriendo, guiando a una persona, deteniéndose ante obstáculos, respondiendo a diferentes comandos, comunicando diferentes cosas e inclusive dirigiendo una orquesta. Es posiblemente uno de los robots humanoides más sofisticados diseñados hasta la fecha. Pero Honda diseñó a Asimo a partir de partes físicas, como plásticos, metales, cables, circuitos, sensores, etc. Pese a ello Asimo es capaz de una impresionante gama de P-funciones. Ahora, considerando a detalle el cuerpo humano, con un cerebro que contiene aproximadamente 10,000 millones de neuronas con aproximadamente 10,000 conexiones sinápticas entre las mismas ¿no es acaso de lo más

plausible que si una persona es un objeto físico, entonces es un objeto físico que tiene una enorme gama de P-funciones correspondientes a su complejidad neurofisiológica? Parece que visto de esa manera el argumento de la función se viene abajo.

Ahora, un problema para el dualista. Si el alma es simple, como habíamos supuesto en el planteamiento sobre identidad personal, no es claro cómo podría pensar, puesto que el pensamiento requiere estructura funcional i.e. partes que ya sea que se especialicen en varias funciones o tales que en ciertas relaciones dan lugar, en combinación con las leyes físico-químicas relevantes, a cierto tipo de función. Tales partes en muchos casos deberían tener una clara contraparte en el funcionamiento del cerebro, por lo que el cerebro basta en este caso en esa explicación. Y es claro que dado que el alma simple no tiene estructura, entonces *a fortiori* no tiene estructura suficiente para el pensamiento y otras actividades funcionales de la mente. Por otro lado, aun cuando el alma fuera compleja, parece que el cerebro es suficiente para caracterizar todos los aspectos funcionales de lo mental, por lo que la navaja de Occam nos obligaría a dispensar del alma como un elemento teórico explicativo en ese dominio. Como bien lo señala Ted Sider en *Riddles of Existence* p. 12:

Pero este argumento es debilitado por la ciencia contemporánea. Los seres humanos han sabido desde hace un tiempo que una parte del cuerpo - el cerebro - está especialmente conectada a la mentalidad. Aún antes del desarrollo de la neurociencia contemporánea era sabido que el daño a la cabeza causaba daño psicológico. Ahora sabemos que trozos particulares del cerebro están conectados con efectos psicológicos particulares. Aunque estamos lejos de poder correlacionar completamente estados psicológicos con estados cerebrales, hemos hecho suficiente progreso que la existencia de esa correlación es una hipótesis razonable. Es sensato concluir que la mentalidad misma reside en el cerebro y que el alma no existe. No es la ciencia cerebral la que demuestra la inexistencia del alma; las almas podrían existir aun si mentes y cerebros estuviesen perfectamente correlacionados. Pero si el cerebro físico explica la mentalidad por su cuenta no hay necesidad de postular al alma.⁴⁸

⁴⁸ "But this argument is undermined by contemporary science. Human beings have long known that one part of the body -the brain- is especially connected to mentality. Even before contemporary neuroscience, head injuries were known to cause psychological damage. We now know how particular bits of the brain are connected with particular psychological effects. Although we're far from being able to completely correlate

Es por ello que el dualismo de sustancias, en tanto explicativo de los aspectos funcionales de la mentalidad, no ofrece suficientes ventajas para ser considerado con seriedad.

IV. El argumento de la concebibilidad

Este argumento no sólo es relevante para el aspecto mental-funcional, sino para el mental-experiencial. Tiene su origen en Descartes, pero la presentación que vamos a explorar tiene como fuente principal otra clase de Shelly Kagan, la clase número 5. Kagan comienza con un experimento mental. Podemos, dice, imaginar una situación en la cual alguien despierta, se levanta, va al espejo y entonces se percata de que su cuerpo es invisible. Esto es claramente imaginable, simplemente pensemos en películas como *The Invisible Man* o personajes de ficción como La Mujer Invisible. De esto a pensar en casos donde no sólo mi cuerpo es invisible, sino que no tengo cuerpo, puedo atravesar paredes, etc; hay un sólo paso. También podemos imaginar que miro hacia abajo y no percibo el resto del cuerpo físico que típicamente asocio conmigo, pero aun así tengo sensaciones; siento como si tuviera un cuerpo pese a que no lo percibo, etc. Para complementar el experimento mental de Kagan, podemos decir que análogamente pueden imaginarse toda clase de escenarios similares. Puedo imaginar, por ejemplo, que estoy levitando por encima del cuerpo físico que típicamente asocio con mi persona (como dicen los charlatanes que hablan de

psychological states with brain states, we have made sufficient progress that the existence of such a correlation is a reasonable hypothesis. It is sensible to conclude that mentality itself resides in the brain, and that the soul does not exist. It's not that brain science disproves the soul; souls could exist even though brains and psychological states are perfectly correlated. But if the physical brain explains mentality on its own, there is no need to postulate souls in addition." *Riddles of Existence*, p. 11.

proyección astral), observándolo desde cierta perspectiva, pese a que yo mismo soy incorpóreo en ese momento. Es decir, es concebible que yo tenga ciertos *qualia* que no estén asociados con cuerpo alguno, como típicamente lo están en la vida diaria. Pero ¿qué podemos inferir de la concebibilidad de estos escenarios?

Creo que podría sostenerse, como hace Kagan, que la facultad de imaginar a una misma cosa (de serlo), como si no fuera única es bastante sorprendente. Podemos considerar casos ordinarios de objetos que conocemos en la vida cotidiana. Por ejemplo, un automóvil. ¿Podríamos imaginar que el automóvil de X tiene y no tiene una ubicación espacial? Parece evidente que no e igualmente hay una lista larga de cosas que no podríamos imaginar. No podríamos imaginar que yo, si soy idéntico al cuerpo físico de Moisés, fuera al cine y no fuera al cine. No podríamos imaginar que nuestros padres existan y no-existan, etc. La razón por la cual sospechamos que no podemos imaginar ninguna de esas cosas es que, dado que son idénticas, el que una cosa tuviera una propiedad que otra no tiene nos parece evidentemente contradictorio. Y de hecho es negar la Ley de Leibniz que, como se ha discutido, no es controversial.

Si este razonamiento es aceptable entonces, dado que aparentemente no podemos imaginar de ningún objeto único que tiene propiedades contradictorias, si imaginar esos escenarios garantiza su posibilidad lógica entonces debe seguirse que la mente y el cuerpo son distintos en tanto tienen diferentes propiedades: esto establecería el dualismo. Sin embargo, hay que decir que es debatible que el concebir que P, implique que P sea posible. De hecho una gran mayoría de filósofos disputa esa idea. Entre los pocos que aceptan eso, o alguna variante, se encuentran David Chalmers en su artículo '*Does Conceivability Entail*

Possibility? y el clásico contemporáneo *The Conscious Mind* y Ludwig Wittgenstein en su *Tractatus Logico-Philosophicus* (TLP, 3.02).

La posición sofisticada de Chalmers la vamos a considerar más abajo, al discutir el argumento del *zombie*. Respecto a Wittgenstein, su posición en el *Tractatus* es diferente de lo que podríamos llamar “la interpretación intuitiva de la noción de concebibilidad”, en la medida en que su teoría de la proposición le hace sostener que toda proposición es bipolar, tiene un polo verdadero, uno falso y es, en último análisis, estructuralmente isomorfa a un hecho posible⁴⁹. La verdad de tal proposición se determina a partir de la relación de isomorfismo entre los signos que la componen y el mundo actual, su falsedad en la ausencia de esta relación. Sin embargo, en tanto Wittgenstein rechaza que las contradicciones y las tautologías sean proposiciones genuinas, debe ser claro que para él cualquier proposición debe ser contingente. Sea esa proposición contingente P: entonces es posible que P y es posible que no-P. El consecuente se sigue del hecho de que la mera posibilidad de P no excluye su necesidad: ‘Necesariamente P \rightarrow Posiblemente P’, es verdadero en cualquier modelo modal donde la relación de acceso entre mundos sea cuando menos serial⁵⁰.

En la medida en que lo de arriba es el caso cualquier proposición tractariana concebible, en tanto proposición, marca una ‘coordenada’ en el espacio lógico y es por ende posible. Pero esta teoría de la proposición no es universalmente aceptada, de manera

⁴⁹ Véase mi artículo publicado *La Crítica de Wittgenstein a la Teoría del Juicio de Russell*, en el libro *Wittgenstein en Español*, 2009.

⁵⁰ Bajo una semántica de Kripke para una lógica modal, un marco modal es una dupla $\langle W, R \rangle$ en donde W es un conjunto no-vacío de “mundos posibles” y R una relación de acceso entre mundos. Decimos que R es serial sii para todo w en W, existe w* en W, tal que w* es accesible a w. Cualquier relación reflexiva es serial.

que no es evidente que concebibilidad implique posibilidad, por lo que de entrada el argumento a lo más muestra que hay evidencia falible de posibilidad.

Sin embargo el materialista tiene una buena respuesta a este caso, que exploraremos más adelante, pues es también una buena respuesta al argumento que veremos a continuación, el argumento del conocimiento.

V. El argumento del conocimiento

El argumento (Jackson, 1997) es como sigue. María es una científica especializada en todo el conocimiento posible de la neurofisiología de la visión de forma que éste conocimiento no se encuentra restringido a los avances tecnológicos o científicos actuales. María sabe todo lo que puede saberse en el campo de la neurofisiología de la visión⁵¹. Sin embargo María está confinada a un cuarto donde sólo existen el negro y el blanco. Es por ello que la científica omnisciente jamás ha percibido color alguno.

Cuando María se libera experimenta entonces aquello que conocía desde el punto de vista objetivo científico sobre los colores. Tiene una representación cualitativa del rojo por primera vez.

¿Qué ocurrirá cuando María sea liberada de su cuarto negro y blanco o le sea dado un televisor de color? ¿Aprenderá algo o no? Simplemente parece obvio que aprenderá algo acerca del mundo y nuestra experiencia visual del mismo. Pero entonces es inescapable que su conocimiento previo estaba incompleto. Pero tenía toda la información física. Ergo, hay más que tener que aquello y el fisicalismo es falso.

Claramente el mismo tipo de argumento del conocimiento podría darse para el gusto, el sonido, las sensaciones táctiles y generalmente hablando para los diversos estados mentales de los que se dice que tienen asociadas una sensación primaria, aspectos fenomenales o *qualia*. La conclusión en cada caso es que los *qualia* son dejados de lado en la narrativa fisicalista. Y la fuerza polémica del argumento del

⁵¹ El argumento tiene varias variantes, a veces María es una experta en física, etc.

conocimiento es que es muy difícil negar la afirmación central de que se puede tener toda la información física sin tener toda la información. (Jackson, 1982, p. 130)⁵²

Según Jackson se sigue que el fisicalismo es falso: Si hay un conocimiento que escapa al conocimiento objetivo científico de la neurofisiología de la visión, entonces hay un conocimiento adicional que se adquiere solamente vía la experiencia. Pero habíamos asumido que María conocía todos los hechos físicos en torno a la visión, si luego no conocía todos los hechos en torno a la visión se sigue que hay hechos no-físicos en torno a ésta.

Entonces, alguna variante de dualismo es verdadero. La razón es que María, pese a ser científicamente omnisciente, se topa con nueva información o nuevos datos que antes no tenía y que además no habría podido tener considerando solamente el punto de vista objetivo neuro-científico/de la física.

Una manera en que Chalmers (2002, p. 250) presenta el argumento de forma sucinta es la siguiente:

- (1) María conoce todos los hechos físicos.
- (2) María no conoce todos los hechos.

⁵² "What will happen when Mary is released from her black and white room or is given a colour television monitor? Will she learn anything or not? It seems just obvious that she will learn something about the world and our visual experience of it. But then it is inescapable that her previous knowledge was incomplete. But she had all the physical information. Ergo there is more to have than that, and Physicalism is false. Clearly the same style of Knowledge argument could be deployed for taste, hearing, the bodily sensations and generally speaking for the various mental states which are said to have (as it is variously put) raw feels, phenomenal features or qualia. The conclusion in each case is that the qualia are left out of the physicalist story. And the polemical strength of the Knowledge argument is that it is so hard to deny the central claim that one can have all the physical information without having all the information there is to have." *The Philosophical Quarterly* Vol. 132, #127, p. 130.

(3) ∴ Los hechos físicos no son todos los hechos.

En el caso de los argumentos discutidos en las secciones IV y de V de este capítulo, parece que el materialista puede valerse de una respuesta. Si esto es así, entonces eso mostraría que en principio es posible evadir el compromiso con el dualismo.

En este caso, una posibilidad es la de adoptar la estrategia fregeana para aclarar el enigma de la identidad. Como es bien sabido Frege introduce en su artículo *Ueber Sinn und Bedeutung* su famosa distinción entre sentido y referencia para poder explicar el enigma de las identidades informativas. El sentido es la intensión de un término o concepto, mientras que la referencia es el objeto del mundo del cual se habla, la extensión. Un ejemplo típico de identidad informativa es del lucero vespertino y el lucero matutino, objetos celestes que se ven en el firmamento a diferentes horas del día. Resulta que el lucero vespertino = lucero matutino = Venus, sin embargo no era posible descubrir *a priori* que aquel objeto al cual se aludía mediante la expresión ‘el lucero matutino’, era el mismo que al que se aludía al usar la expresión, ‘el lucero vespertino’. De forma que cuando finalmente la astronomía estableció que se trataba del mismo objeto, sólo entonces se supo la verdad de ‘el lucero matutino = el lucero vespertino’. Algo metafísicamente necesario, según los creyentes en la necesidad de la identidad⁵³, era epistémicamente incierto.

En el caso de Venus parece que podríamos haber imaginado que no hubiera un lucero matutino y si uno vespertino o viceversa. Pero en ello evidentemente habríamos estado equivocados, en la medida en que al ser el mismo objeto no puede ser que Venus exista y no exista. Lo que pasa con lo que imaginamos, en ese escenario, es que por razones

⁵³ Véase Kripke, 1971.

de nuestra ignorancia lo describimos erróneamente. Así, un escenario en donde imaginemos que no hubiera habido lucero matutino pero si vespertino simplemente sería un escenario en donde Venus no se hubiera visto en el firmamento en las mañanas.

Podemos explicar que la verdad de un enunciado identidad sea conocida *a posteriori* empleando algunos recursos de la filosofía del lenguaje de Frege. Decimos en este caso que las expresiones ‘lucero matutino’ y ‘lucero vespertino’ tienen la misma referencia, pero diferente sentido. El sentido es alguna descripción que se asocia con el término, en el caso del ‘lucero matutino’ sería ‘aquel objeto celeste que se ve por las mañanas’, etc. Y análogamente con el ‘lucero vespertino’. Mientras que trivialmente sabemos que ‘el lucero vespertino = el lucero vespertino’ en la medida en que ambas expresiones tienen el mismo sentido, no podemos saber eso en casos donde el sentido es diferente. El sentido en este caso es un ‘modo de presentación’, algún objeto se me presenta bajo el modo de presentación de ser lo que satisface cierta descripción. Así el materialista puede decir que lo que pasa en el caso de María no es que se entere de nuevos hechos, sino que conoce viejos hechos, los hechos sobre la neurofisiología del color, bajo un nuevo modo de presentación.

Respecto al argumento de concebibilidad en IV donde aparentemente podía imaginar a mi mente y mi cuerpo como entidades independientes, me parece que estamos en posición de decir algo. Habíamos supuesto que el que algo pudiese concebirse como si no fuese único (bajo distintos modos de presentación) entonces ello no era una sólo objeto, sino objetos lógicamente independientes. Sin embargo parece que en esta sección se ha dado un ejemplo de un caso en donde parecía concebible que un mismo objeto (Venus),

bajo diferentes modos de presentación, no existiera en uno y si existiera en otro, pero se ha mostrado que realmente no era posible. En este caso era sólo una ilusión y ello nos da razones para rechazar el supuesto fuerte detrás del argumento de concebibilidad de que: si puede imaginarse alguna diferencia entre x e y , entonces son diferentes. Quizá lo que pasa es que al conocer un mismo evento físico, bajo diferentes modos de presentación, se asume erróneamente y de manera precipitada que podría haber una diferencia *de re* y no sólo *de dicto* i.e. una diferencia en los objetos mismos y no sólo en las proposiciones que usamos para referirnos a ellos.

En principio no parece que este argumento sea defectuoso, por lo que al menos algunos casos en dónde nos sentimos inclinados a atribuirle diferencias a objetos son realmente casos en donde hay diferencias a nivel proposicional, falta de información, etc. Plausiblemente descripciones del mismo fenómeno a diferentes niveles de abstracción; desde diferentes ciencias especiales, etc; que de entrada parecieran indicar diferentes fenómenos son posibles y en esa medida esos casos no serían suficientes para establecer el dualismo⁵⁴.

Pero este argumento no muestra que es plausible sostener una posición eliminativista respecto a los *qualia*. De conceder el fisicalismo este argumento mostraría, pienso, que en la medida en que percibir un objeto bajo el modo de presentación cualitativo es algo que efectivamente hacemos para una mayoría de objetos, es claro que esas

⁵⁴ Supongamos que tenemos dos descripciones de un sujeto: a nivel de DNA y a nivel sociocultural (e.g. se llama X, trabaja en Y, vive en Z, etc.) Desde luego es posible tener ambas a la mano y no saber que hacen referencia al mismo objeto sin que de ello se siga que existe una sustancia no-física de la socio-culturalidad, etc.

propiedades son reales. Aun si es posible referir a un objeto bajo un modo de presentación no cualitativo, es claro que la evidencia de la existencia de cualidades es de primera persona y que, por ende, cualquier fisicalismo efectivo debe acomodarlas. Pero el fisicalismo y funcionalismo ortodoxos no tienen, en tanto eliminativistas, una buena forma de acomodar a las cualidades o de explicar cómo puede ser que lo material tenga un aspecto cualitativo.

VI. Dualismo y Experiencia (*Qualia*)

Parece que si lográbamos vincular al alma poderosamente con la identidad personal, quizá podríamos haberla empleado para vincularla a los *qualia* y matar dos pájaros de un tiro. Pero tampoco parece que un objeto metafísicamente simple podría explicar coherentemente la complejidad de los aspectos cualitativos de lo mental. Es evidente que los *qualia*, por decir, en el campo visual, tienen partes, por ejemplo el color azul del cielo que ocupa mi campo visual, está extendido en él, tiene parte inferior y superior, etc. En el campo auditivo, también, cuando pensamos en algún tono extendido, etc. Por otro lado, si el alma fuera mereológicamente compleja esta objeción no puede plantearse. Empero, al ser una sustancia radicalmente diferente, uno podría preguntarse en qué consiste su complejidad. ¿Podría ser espiritual y ocupar el espacio-tiempo de forma que tenga estructura mereológica? Si lo ocupa, ¿en dónde reside? Quizá el alma es del tamaño del cuerpo y ocupa exactamente la misma región espacial en t que su cuerpo correspondiente. Por otro lado, quizá el alma está flotando por el universo. Si lo segundo es el caso, más vale que no esté demasiado lejos, puesto que el tope de velocidad de la información en nuestro mundo, es la velocidad de la luz (aproximadamente 300,000 Km/s). Ciertamente si mi alma

experimenta el azul cuando fotones de azul impactan mi retina, más vale que su reacción sea simultánea con el impacto (o lo suficientemente cercana temporalmente para que su existencia no pueda descartarse empíricamente).

Al menos dos objeciones pueden plantearse a estas especulaciones. Éstas se siguen del siguiente disyunto: ya sea que hay interacción cuerpo-alma o no la hay. Si la hay, la interacción es completamente misteriosa. ¿Cómo pueden interactuar dos sustancias radicalmente diferentes? Comprendemos bien, gracias a la física, cómo hay interacción material, pero el alma no es material, por lo que es totalmente misterioso cómo algo así interactúa con el cuerpo.

Supongamos que el alma no interactúa con el cuerpo. Esta opción es sumamente implausible en la medida en que tenemos muy buena evidencia de correlaciones entre sucesos corporales y sucesos mentales (*qualia*). Si pico mi dedo con un alfiler, experimento de inmediato un dolor en la punta, etc. Pero además, parece claro que toda la información cualitativa y funcional presente a mi mente corresponde con lo que ocurre en el mundo físico, en tanto es posible su manipulación. Sabemos que el cuerpo está biológicamente capacitado para recabar información de sus entornos, pero si el cuerpo no interactúa con el alma es milagroso que ésta perciba con casi perfecta claridad el medio en que está el cuerpo y sus interacciones con él.

Pero además, tenemos buenas razones para suponer que no ocurre tal tipo de interacción. Esto se debe a que en física tenemos un andamiaje conceptual de enorme poder predictivo el cual no involucra ninguna otra cosa excepto las entidades postuladas en nuestras teorías físicas. A la tesis de que sólo lo físico interactúa con lo físico se le llama

‘cerradura causal del mundo físico’⁵⁵, y es una tesis masivamente respaldada por la evidencia empírica. Si esto es verdad, entonces no hay interacción, lo cual es afortunado, puesto que de haberla parecería imposible tener una explicación de la misma. Pero entonces parece que hay buenas razones para aceptar la cerradura causal del mundo físico y por ende, para rechazar la interacción cuya dirección es únicamente de alma a cuerpo.

Si no hay interacción del alma con el cuerpo entonces hay al menos dos posibilidades: la posibilidad de que el cuerpo y el alma nunca interactúen y la posibilidad de que, pese a que el alma no interactúe con el cuerpo, éste sí con el alma. Ésta última teoría es compatible con una interpretación más débil de la tesis de la ‘cerradura causal del mundo físico’. Fuertemente interpretada, la tesis nos dice que la única interacción es a nivel físico, pero débilmente interpretada, lo que nos dice es que sólo la interacción de lo físico es relevante para explicar lo físico. Esto último es consistente con que lo físico afecte lo mental, aun si ello no juega ningún rol explicativo en la interacción de lo físico. Claramente eso corresponde a la teoría de la interacción sólo en la dirección cuerpo-alma.

Por el lado de la teoría de la no interacción, ésta parece aún más increíble que la teoría anterior, por las razones que ya se han venido dando. La mejor teoría de no interacción es la de la armonía preestablecida de Leibniz⁵⁶, o alguna variante de la misma. Bajo esta posición cuerpo y alma jamás interactúan, pero sus percepciones están coordinadas a la perfección de forma que cuando mi cuerpo es golpeado mi alma

⁵⁵ Véase Vicente, 2006.

⁵⁶ Véase el *A Critical Exposition of the Philosophy of Leibniz*, de Bertrand Russell.

experimenta dolor y cuando mi alma tiene la voluntad de mover mi cuerpo, éste se mueve. Cada uno en sus respectivos dominios nómicos.

Esta teoría es claramente increíble, si bien, *prima facie* consistente. Un claro problema es explicar la sincronía de las sensaciones. Explicaciones que funcionan muy bien con órganos específicos de nuestro cuerpo, como la teoría evolutiva con selección natural, son claramente incapaces de explicar nada respecto a esta sincronía. La evolución actúa mediante las leyes de la bioquímica sobre organismos biológicos⁵⁷, cosa que no es el alma. Es claro por qué la evolución beneficiaría el desarrollo de ciertas capacidades perceptuales, pero no haría nada para explicar porque deben corresponder *qualia* a ese tipo de capacidades en la medida en que aun sin *qualia* el cuerpo se comportaría exactamente de la misma manera: recuérdese que esta teoría presupone la no-interacción. Esto implica arrojar por la borda una de nuestras mejores herramientas para explicar el desarrollo de organismos en el mundo natural.

Pero no sólo eso, sin traer a Dios a la ecuación parece que es imposible explicar en principio tal sincronía, pero ésta es una carga teórica adicional de la teoría del alma con armonía preestablecida. Bajo esta concepción, es puramente un capricho del mundo que exista esta sincronía, no hay una razón suficiente para ella. No hay nada en las leyes de la física que pudiera explicarnos la sincronía, pues la física trata sólo con objetos físicos y sus interacciones. Alguien puede aceptar que haya hechos brutos en ciertos dominios explicativos, pero la sincronía entre mente y materia no es uno de ellos. La interacción (alma) mente-materia tiene un enorme poder explicativo en nuestra vida diaria y para

⁵⁷ Véase Dennett, 1995.

explicar todo tipo de fenómenos respecto a cómo nos relacionamos con el mundo. Abandonar la interacción implicaría renunciar a todas esas explicaciones plausibles. Seguro que debe haber teorías que preserven esas explicaciones y tengan mayor poder explicativo que ésta.

Si es falso tanto que no existe interacción, como que ésta sólo va del alma al cuerpo⁵⁸, pero verdadero que existe el alma, entonces el epifenomenalismo es verdadero. Ésta es la tesis de que la dirección causal sólo va de un lado, de lo físico a lo mental (el alma) pero no al revés. Esto implicaría, particularmente, que cuando yo deseo tomar agua y siento que mi volición produce en el mundo el efecto de que mi brazo toma un vaso de agua, soy víctima de una ilusión. No tengo, de facto, control de mis acciones, aunque parezca que mi volición mueve mi cuerpo esto es sólo un espejismo. Lo que le ocurre a mi cuerpo afecta mi alma y por lo tanto mis *qualia* (en esta teoría), pero no a la inversa. Seguro que una teoría así es implausible. Pienso en el experimento mental de Smullyan en *An Unfortunate Dualist* que involucra a un sujeto que desea suicidarse pero sin ir contra sus creencias religiosas o afectar a sus seres queridos: el hombre adquiere un medicamento que sólo elimina la conciencia (los *qualia*), sin embargo dado el epifenomenalismo, el cuerpo del sujeto continua comportándose exactamente de la misma forma después de tomarla, comentando inclusive: “maldición, esta cosa no sirve, es obvio que aún tengo un alma y sigo sufriendo tanto como siempre” (Chalmers, 2002, p. 31). Por ésta y por las razones que expuse más arriba, creo que debemos abandonar este enfoque.

⁵⁸ Quizá algo así podría tener coherencia bajo filosofías como las de Berkeley o Malebranche, sin embargo dada la tesis de la cerradura causal del mundo físico, que aquí aceptamos como empíricamente respaldada, no consideraré esta posibilidad en este trabajo.

Hemos visto entonces que la teoría del alma implica que hay o no interacción de ésta con la materia, pero ninguno de los disyuntos es plausible. Si ambos son falsos, como parece hay razones para creer, la teoría del alma compleja es falsa, puesto que, por razones de lógica elemental, lo que implica lo falso es falso.

Una tercera objeción, de corte metodológico, integra y sintetiza bien la discusión hasta ahora. La objeción consiste en notar que aun cuando el problema de los *qualia* es sumamente difícil de tratar, la postulación del alma no nos ayuda en nada a resolverlo. Lo único que hace es agregar una entidad *ad hoc* en donde puedan estar esas cualidades. Pero esto no nos ayuda a explicar cómo podrían estar ahí, ni como surgen. Es decir, no se ve por qué la existencia del alma tiene poder explicativo en torno a los *qualia*. Puesto que ya sea que estos dependen ontológicamente del alma, un objeto en sí mismo misterioso, o son primitivos. Pero lo primero es tan implausible como el materialismo con superveniencia (no se diga el materialismo reduccionista), en tanto no se ve cómo es que la existencia del alma garantiza la existencia del *qualia*. No hay ninguna relación conceptual entre alma y *qualia* que la garantice y mucho menos una relación lógica, por otro lado la existencia de una relación nomológica requeriría la postulación de éste dominio extra de entidades y leyes, la cual es implausible por las razones que dimos más arriba⁵⁹. Por todo ello es que no es claro cómo podría haber una dependencia ontológica.

⁵⁹ Sin embargo ésta es quizá una de las alternativas prometedoras para el dualismo, aunque no el de sustancias. Chalmers la explora a detalle en *The Conscious Mind*. Un resultado de esa exploración es el deshacerse del intermediario i.e. el alma o mente: las leyes puente conectan los dominios nómicos de lo material-funcional con el dominio de las experiencias dando lugar a un dualismo de propiedades o funcionalismo no-reduccionista.

En síntesis, los *qualia* no supervienen de las almas. Esto quiere decir que podría haber variaciones en experiencia sin variaciones en las almas. ¡Un mundo de almas zombie es lógicamente posible! Por otro lado, si los *qualia* son ontológicamente primitivos, entonces postular el alma para explicar los *qualia* es completamente redundante.

Ya vimos, en todo caso, que la teoría del alma compleja es sumamente implausible por otras razones y seguro esto debe motivarnos en pensar en otras soluciones al problema de los *qualia*. Unas que mejor integren lo que sabemos del mundo físico y del cerebro con las cualidades subjetivas de la experiencia.

En este punto creo que es razonable concluir que el dualismo de sustancias no es teóricamente útil para resolver ni el problema de la identidad personal, el problema del funcionamiento de la mente y *prima facie* tampoco el problema de la experiencia.

CAPÍTULO II

PROBLEMAS PARA EL FISCALISMO REDUCCIONISTA

En el capítulo anterior se exploraron diversos argumentos a favor del dualismo y se vio que muchos de ellos distaban de establecer algo suficientemente poderoso como para reconsiderar abandonar el fisicalismo. Sin embargo el fisicalismo reduccionista tiene la dificultad de negar en muchos casos la indubitabilidad de los *qualia*, o acepta los *qualia* y los trata teóricamente como otra cosa: me viene a la mente la sugerencia de algunos funcionalistas de tratar a un *quale* solamente como lo que ocupa un nodo causal-funcional en alguna estructura⁶⁰. Dado que considero que el reduccionismo respecto a los *qualia* es inaceptable, como discutiré más abajo, entonces me hubiese gustado que hubiera habido una mayor cantidad de argumentos fuertes en favor del dualismo.

En la siguiente sección voy a retomar la argumentación contra el fisicalismo, enfocándome específicamente en el problema que representan los *qualia* para toda teoría fisicalista. Pienso que en lo hasta ahora discutido se ve que no hay razones para suponer que los aspectos mentales-funcionales de la mente requieren de más que alguna estructura material sofisticada que pueda realizar una estructura causal-funcional, como el cerebro, el cual emerge por procesos evolutivos. Replantearé el argumento de Jackson, uno de los argumentos típicos en contra del materialismo, y esbozaré dos nuevos argumentos que, hay razones para suponer, toda teoría correcta de la mente debe acomodar.

⁶⁰ Véase Lockwood, 1989, Cap. 3 “Functionalism”.

En el siguiente capítulo mostraré que una teoría de corte Russelliano los acomoda y que tenemos razones para suponer que esta teoría no adolece de los defectos típicos del dualismo y fisicalismo ortodoxos.

I. Argumentos en contra del fisicalismo

Voy a replantear los principales argumentos en contra del fisicalismo, ahora con un objetivo diferente. En el capítulo anterior habíamos visto que al menos los aspectos funcionales de lo mental eran en principio explicables de forma fisicalista, pero nos mantuvimos escépticos respecto a los aspectos fenoménicos de lo mental. Lo que voy a hacer a continuación es mostrar cómo esos mismos argumentos constituyen buena evidencia de que los aspectos fenoménicos de lo mental simplemente no son explicables desde un fisicalismo ortodoxo, sin importar cuanta estructura podamos añadir a la teoría fisicalista última.

Una característica de estos argumentos es que la experiencia va a jugar un rol central en los mismos. Estos son probablemente los más discutidos en la literatura contemporánea sobre el tema y si puede mostrarse que una perspectiva Russelliana arroja luz sobre ya sea los errores ó aciertos de los mismos entonces podemos decir que estamos alcanzando una perspectiva que, en la medida en que incorpora estos elementos, cuenta con mayor poder explicativo.

Entre los argumentos más prominentes que tratan de mostrar la falsedad del fisicalismo se encuentran los siguientes:

El argumento del conocimiento (Jackson, 1997):

Este argumento ya se había expuesto. Aun así, la solución Russelliana debe ser capaz de confrontarlo.

María está atrapada en un cuarto en blanco y negro, sin embargo ahí aprende todo lo que puede saberse sobre la naturaleza física del mundo. Aun así, cuando sale del cuarto y ve el color rojo por primera vez ha aprendido un nuevo dato, a saber ‘cómo se ve el rojo’. Pero dado que sabía todos los datos físicos, entonces el fisicalismo es falso i.e. los datos físicos no son todos los datos⁶¹.

El argumento del zombie (Chalmers, 1996):

Un duplicado físicamente idéntico de mí, pero que no posea *qualia* i.e. un zombie filosófico, es concebible. La concebibilidad implica posibilidad lógica. Pero entonces si un duplicado físicamente idéntico a mí, pero que no posea *qualia*, es concebible entonces es posible. Por lo tanto el fisicalismo es falso puesto que si fuese verdadero entonces todo duplicado físico de mí, tendría que ser un duplicado de mí *simpliciter* (*qualia* incluido).

Este argumento es una variante del argumento de los *qualia* ausentes, avanzado para refutar algunas variantes del funcionalismo.

⁶¹ La noción de “dato” corre el riesgo de ser interpretada epistémicamente, pero lo que queremos decir es más fuerte. Que haya datos no-físicos implicaría que hay un dominio de hechos no-físicos acerca de los cuales los datos físicos por sí solos no nos dicen nada, ausentes ciertas leyes puente.

El argumento modal (Kripke, 1980):

De acuerdo a Kripke⁶², no hay identidades contingentes. Sin embargo, aquellos que han sostenido que alguna sensación es idéntica a un estado cerebral han concedido que la identificación en este caso es contingente. Bajo este esquema, la explicación usual es que lo contingente es el modo de presentación. Así, en el caso de identidades como 'H₂O = Agua' se dice que lo contingente es que el H₂O sea la sustancia que identificamos en la medida en que es transparente, insípida, líquida en cierto rango de temperatura, etc. Pero concedemos, por el argumento de la necesidad de la identidad, que necesariamente H₂O es agua. Sin embargo, en el caso de una sensación, como el dolor, el modo de presentación no es otra cosa que el dolor mismo. El dolor simplemente **es** la sensación que siento cuando digo que me duele. Pero entonces la identidad entre las sensaciones y los estados cerebrales tendría que ser necesaria. Sin embargo parece evidente que seres fisiológica o materialmente distintos a nosotros, como marcianos o robots, podrían sentir dolor. Pero si esos seres pueden sentir dolor, entonces es posible que el dolor sea diferente a cierto estado cerebral, contradiciendo nuestro supuesto. Por lo tanto, el fisicalismo es falso.

Hay que tomar en cuenta que el golpe principal que asesta el argumento de Kripke es contra teorías de la identidad eliminativistas. La razón es que estas teorías sostienen simplemente que un evento mental es un evento cerebral, pero que hablar de cualidades de la experiencia simplemente es hablar de un modo de presentación contingente de lo mental, análogamente a como hablar de calor es hablar de un modo de presentación contingente del movimiento molecular. Dado que es concebible que nuestros cuerpos hubieran sido tales

⁶² Véase Kripke, 1971.

que, al aumentar el movimiento molecular, hubiesen sentido la sensación de frío. En este caso, la sensación de calor no es idéntica al calor como movimiento molecular, sino sólo es la forma en cómo nosotros aprehendemos el calor. Sin embargo no podemos decir lo mismo en el caso de experiencias como el dolor, dado que en ese caso el modo de presentación simplemente colapsa en el referente. De manera que decir que el verdadero dolor, como estado cerebral, pudo haberse sentido diferente es decir algo absurdo.

Sobre estos tres argumentos John Perry (Perry, 1999) ha sostenido que en la medida en que muestran con mayor fuerza el problema de la experiencia para el fisicalismo entonces, plausiblemente, cualquier teoría de la mente fisicalista adecuada debe ofrecer respuestas a los mismos.

II. Algunas consideraciones sobre el problema mente-cuerpo

El debate sobre el problema mente-cuerpo ha sacado a la luz varios supuestos importantes, todos ellos plausibles, que no pueden ser todos verdaderos bajo la concepción ordinaria de la física y el supuesto de que ésta es la ciencia más metafísicamente fundamental. Sin embargo, parece razonable exigir de cualquier teoría que pretenda explicar qué es la mente el que pueda ofrecer un diagnóstico sobre qué es lo que pasa con cada uno de estos supuestos.

i) Cerradura Causal del Mundo Físico (CCMF): Este es un supuesto metodológico que permea la más avanzada de las ciencias duras no estrictamente formales i.e., la Física. En esencia, todo modelo de la física presupone que eventos físicos causan y a su vez son causados por eventos físicos solamente. Así, toda cadena causal, al interior de la física, es una cadena causal que comienza y termina con eventos físicos y en donde se considera que,

para cualesquiera eventos físicos $e1$ y $e2$, tales que $e1$ causa $e2$, no es necesario el tomar en cuenta ningún evento no físico para explicarse el que $e2$ haya ocurrido. La evidencia a favor de este supuesto es abundante. Crucialmente la física es la ciencia más exitosa con la que contamos en términos de explicación y predicción, éxito que se ha logrado sin postular fuerzas diferentes a las cuatro interacciones fundamentales: gravedad, nuclear débil, nuclear fuerte y electromagnetismo.

ii) **Influencia de lo Mental (IM):** Este supuesto recibe considerable apoyo de nuestra visión intuitiva de las cosas. Usualmente consideramos que el tener la intención de hacer algo, más la volición de hacerlo, conducen directamente a la acción. Por citar un ejemplo, consideramos que si deseamos salir a correr por las mañanas, entonces podemos, mediante el uso de nuestra voluntad, mover a nuestro cuerpo para vestirlo y salir a trotar. Podemos tomar una lata de refresco si así lo queremos; expresar nuestras ideas y en general interactuar con el mundo. Esta es una influencia de lo mental a lo físico y definitivamente parece recibir soporte directo de nuestras sensaciones de poder afectar el mundo físico a partir de nuestros deseos y voliciones.

Indubitabilidad de los Qualia (IQ): Una cosa de la que una cantidad considerable de filósofos (y plausiblemente de no-filósofos) están seguros es de que tienen ciertas cualidades fenomenales (qualia) y estas son conocidas de manera directa y sin inferencia. Para una mayoría cualquier solución al problema mente-cuerpo tiene que tomar en cuenta la indubitabilidad de los qualia. (Véase Chalmers, 1996)

Unidad de la Conciencia (UC): Solemos aceptar que la conciencia tiene una propiedad característica, esto es, la de sentirse como un todo unificado. Mediante la introspección la

persona promedio no será capaz de detectar alguna división en el campo visual. Es evidente también que nuestras sensaciones olfativas, gustativas, visuales, táctiles y auditivas pueden todas experimentarse simultáneamente.

Estos cuatro supuestos permean nuestra manera de pensar en el problema mente-cuerpo y si bien puede decirse que no son los únicos supuestos, quizá podemos contarlos entre los más importantes. Más adelante exploraré como la teoría de Russell aborda cada uno de ellos.

III. Algunos argumentos anti-fisicalistas explorados a fondo.

En esta sección voy a ver con más cuidado el argumento del zombie y un trilema que surge de considerar algunos de los supuestos de la sección VIII.

A) Un trilema incómodo

El trilema al que me refiero surge de rechazar algunos de los primeros tres supuestos. Es incómodo en la medida en que todas sus conclusiones son contraintuitivas. Si se abandona IQ la postura colapsa en una versión cruda de fisicalismo que sostiene que los qualia son una ilusión. Si se abandona IM la postura colapsa en un epifenomenalismo i.e. el abandono de la noción de que lo mental influye en lo físico. Si se abandona CCMF la postura colapsa en el dualismo interaccionista, y tenemos (cuando menos) dos sustancias radicalmente diferentes de las cuales no podemos explicar cómo interactúan:

- a) Si somos fisicalistas reduccionistas entonces tenemos que aceptar que lo mental (como sustancia independiente) es una ilusión y no es más que lo material, ya sea ciertos estados cerebrales o estados funcionales, pero si somos dualistas

interaccionistas entonces tenemos que aceptar que, contrariamente a la ortodoxia científica, la física no está causalmente cerrada. Sin embargo, si somos epifenomenalistas entonces tenemos que aceptar que si bien existe lo mental, esto no provoca ningún efecto físico.

b) Pero algunos filósofos (ej: John Perry) han argumentado convincentemente que entre las muchas alternativas en la filosofía de la mente, éstas básicamente son de corte materialista o de corte dualista o de corte epifenomenalista.

∴ Por lo tanto (si estos filósofos tienen razón y la discusión tiene sentido), o la mente en sus aspectos cualitativos (como sustancia independiente) es una ilusión o la física no está causalmente cerrada o lo mental no interactúa con lo físico.

Todas las conclusiones resultan indeseables en tanto contraintuitivas y en la medida en que nos obligan a renunciar a uno de nuestros supuestos sin, al menos de entrada, permitir una explicación satisfactoria sobre porque ese supuesto, pese a su falsedad, era generalmente considerado como verdadero. El abandono de CCMF introduce entidades no contempladas en la física, que nos obligan a llevar a cabo una especie de revisionismo científico y que además nos deja en las sombras respecto a la naturaleza de estas entidades y su habilidad para interactuar con lo físico. El abandono de IQ es altamente implausible (pace Dennett, Churchland, et al). El abandono de IM, si bien no entra en conflicto con la física, parece totalmente contrario a lo que creemos se da todo el tiempo en nuestras interacciones con objetos macroscópicos.

B) Dos versiones del argumento del zombie.

He parafraseado el argumento del *zombie* de Chalmers en dos versiones que me permitirán explorar más adelante qué es lo que puede decirse respecto al mismo. Chalmers no presenta su argumento de ninguna de las dos maneras.

- i) Concebibilidad implica posibilidad.
- ii) Lo mental-funcional y lo mental-experiencial son categorías diferentes. La primera concierne la estructura psicológica/cognitiva del sujeto (o sujeto + medio ambiente), de tal que, dada cierta estructura a ciertos estímulos se producen ciertas respuestas. Lo mental-experiencial tiene que ver con ‘como se siente ser algún ser’, referente a cómo son las experiencias visuales, gustativas, auditivas, etc y las emociones de tal ser.
- iii) Un zombie es un duplicado de una persona en todos sus aspectos físicos y mental-funcionales pero no en sus aspectos mental-experienciales.
- iv) Los zombies son concebibles.
- v) Si los zombies son posibles entonces el materialismo, la tesis de que todo lo que hay en el mundo es solamente material, es falso.

-
- i) ∴ Los zombies son posibles. i) y iv)
 - ii) ∴ El fisicalismo es falso. De Conclusión i) y v)

Las distinciones trazadas en la premisa ii) son cruciales para que el argumento sea válido. Originalmente, una versión del argumento conocida como el argumento del *qualia* ausente se había empleado en contra de ciertas variantes fuertes de funcionalismo. El funcionalismo, como ya dijimos más arriba, es una postura fisicalista bajo la cual lo relevante a lo mental es la estructura causal-funcional (Lockwood, 1989). En esta postura algún evento es mental en la medida en que ocupa cierto nicho en la estructura causal-funcional.

Un ejemplo plausible de un ser con una estructura causal-funcional es un robot, pensemos, por ejemplo, en la sonda *pathfinder* que exploró la superficie marciana en 1996. El *pathfinder* tiene una serie de sensores y un hardware que corre varios programas, es evidente que la sonda es capaz de recibir estímulos externos y responder a los mismos. La estructura causal-funcional de la sonda es lo que le permite hacer esto. Diríamos, por ejemplo, de la sonda *pathfinder* que el percibir una piedra marciana para la misma simplemente es que haya un estímulo visual en uno de sus sensores que sea interpretado de alguna manera por el algoritmo que se instancia de manera concreta en su *hardware*.

Así, la idea del argumento es que podemos concebir a entidades como el *pathfinder*, del cual pocos dirían que tiene cualidades fenomenales. Si podemos concebir que el *pathfinder* tenga una estructura causal-funcional, i.e. un esquema mental-funcional, pero no mental-experiencial (*qualia*) entonces ¿no es acaso concebible algún ser que nos duplique en nuestras características físicas y mental-funcionales pero no mental-experienciales? Nótese que la única diferencia que pudiera tener alguna relevancia científica aquí sería el material del que está hecha la sonda en contraste con el material biológico del cerebro

humano. Pero es de lo más extraño suponer que estar compuesto por carbono sea una condición necesaria para los *qualia*. En parte la lección de la literatura sobre realizabilidad múltiple es que diferentes tipos de materiales puede instanciar la misma propiedad (mental o funcional) siempre y cuando tengan una estructura isomorfa o cuando menos un *embedding*. Si es concebible y concebibilidad implica posibilidad, tenemos las conclusiones señaladas.

Otra paráfrasis del argumento del zombie es la siguiente:

- i) M es la conjunción de todas las proposiciones sobre hechos mentales (*qualia*).
- ii) F & L es la conjunción de todas las proposiciones sobre hechos físicos y las leyes de la física.
- iii) Una proposición ϕ es consecuencia lógica de un conjunto de premisas Σ si y sólo si todo modelo de Σ es modelo de ϕ .

Pero Russell (1927a, 1927b, 1948) ha argumentado que la física sólo nos proporciona conocimiento estructural, es decir, que la física nos dice que el mundo tiene cierta estructura pero es silenciosa respecto a aquello que instancia la estructura. Elaborando, la física nos informa sobre las relaciones que se dan entre los objetos del dominio de entidades físicas, pero no dice nada que nos permita individuar esos objetos relativamente a otra estructura abstracta isomorfa a la primera. Pero un *quale* definitivamente es capaz de ser individuado relativo a cualquier estructura isomorfa a una estructura de *qualia* pero con diferentes objetos ocupando el mismo nicho. Es individuado en la medida en que hay algo que es tener tal *quale*. En conclusión la física no nos dice nada sobre los *qualia* en sus aspectos intrínsecos, algo que discutiré a fondo más abajo.

Esto es aceptado por varios, entre ellos Chalmers (1996), Lockwood (1989), Maxwell, (1970), etc. Por lo tanto:

iv) Hay un modelo de (F & L) que no es modelo de M.

∴ Las proposiciones sobre hechos y leyes de la física no tienen como consecuencia lógica las proposiciones sobre hechos mentales cualitativos. De iii) y iv).

Pienso que estos argumentos tienen suficiente fuerza para mostrar cuando menos que hay un hueco explicativo entre nuestra concepción de la materia y la existencia de la experiencia. Y que este hueco explicativo no puede superarse razonando demostrativamente de las verdades de la física, entendidas a grandes rasgos, a las verdades de la experiencia. Si esto es así, entonces si bien no es un golpe definitivo contra las tesis fisicalistas sí que disminuye el rango de las alternativas fisicalistas posibles.

Si todo lo que he venido diciendo se sostiene pienso que podemos mantener que una teoría fisicalista aceptable tendría que aceptar la experiencia como ineliminable y no fundamentada en las verdades de la física pero también compatible con las mismas; sin postular propiedades o sustancias adicionales que la hicieran colapsar en un dualismo. Veremos que la posición Russelliana tiene esos elementos.

CAPÍTULO III
EL PROBLEMA DIFÍCIL DE LA MATERIA: LA TEORÍA DE
BERTRAND RUSSELL⁶³

En síntesis, la teoría de Russell es una teoría no reduccionista. Algunos exégetas sostienen que es una teoría de identidad mente-cerebro no reduccionista, sin embargo creo que hay interpretaciones de su teoría que no la convierten en una teoría de identidad. Esta discusión la iremos elaborando sobre la marcha.

De entrada podemos decir que Russell acepta IQ, pero al mismo tiempo acepta CCMF e IM. Russell logra esto en la medida en que, para él, aquello que conocemos de manera directa y con certeza viene por el lado de lo mental. Russell considera, al contrario de muchos filósofos contemporáneos, que la materia es misteriosa excepto donde la materia resulta ser nuestro cerebro, el cual conocemos de manera directa. Lockwood comenta:

De lo que Russell se percató, y de lo cual pocos otros filósofos lo han hecho, es que las objeciones al materialismo están implícitamente basadas en el supuesto de que estamos, después de todo, en contacto directo a través de la percepción con la naturaleza intrínseca o inherente de la realidad material y, estando así familiarizados con la misma, nos percatábamos que era radicalmente diferente de nuestras vidas mentales. (1989 p. 158)⁶⁴

Es decir, para Russell todo el problema mente-cuerpo está, de entrada, concebido erróneamente en la medida en que la mayoría de quienes lo investigan creen estar

⁶³ Obviamente ésta es una referencia, espero cómica, al así llamado problema difícil de la conciencia. Término introducido por Chalmers.

⁶⁴ "What Russell realized, and very few other philosophers have, is that these objections to materialism are implicitly predicated on the assumption that we are, after all, transparently acquainted in perception with the inherent or intrinsic nature of material reality – and, being so acquainted, see it to be radically different from the reality of our own mental lives." *Mind, Brain and the Quantum* p. 158.

familiarizados con la naturaleza intrínseca de la materia. Russell piensa que, por el contrario, no sabemos exactamente qué sea la materia.

A lo largo de varios libros, comenzando por *The Problems of Philosophy* y pasando por *Introduction to Mathematical Philosophy*, Russell avanzó una serie de argumentos a favor de la tesis de que del mundo físico sólo tenemos conocimiento estructural, con la excepción de una región del mismo.

Antes de entrar en detalle en los pormenores de la postura de Russell será conveniente ver cómo algunas dudas, como las que sugiere el experimento mental del molino de Leibniz, son de inmediato bloqueadas en el marco de la teoría russelliana. Este es el argumento de Leibniz:

Debe confesarse que la Percepción, y aquello que depende de ella, son inexplicables por causas mecánicas, esto es, por formas y movimientos. Suponiendo que hubiese una máquina cuya estructura produjese el pensamiento, la sensación y percepción, podríamos concebir la misma como aumentada en tamaño con las mismas proporciones hasta que uno pudiera entrar en su interior, como lo haría dentro de un molino. Al entrar a ella sólo encontraría piezas trabajando una sobre la otra, pero nunca encontraría nada que explicara la percepción. (Leibniz, 1974 p. 457)⁶⁵

La réplica inmediata de alguien que sostenga la postura de Russell sería que aquello que Leibniz considera, en el experimento del molino, como su impresión de la materia sólo es algún percepto en su mente, producto final de una cadena causal cuyo origen está en el molino. Por lo mismo Leibniz, o quien quiera que entre al molino, no está en posición de

⁶⁵ "It must be confessed, however, that Perception, and that which depends upon it, are inexplicable by mechanical causes, that is to say, by figures and motions. Supposing that there were a machine whose structure produced thought, sensation, and perception, we could conceive of it as increased in size with the same proportions until one was able to enter into its interior, as he would into a mill. Now, on going into it he would find only pieces working upon one another, but never would he find anything to explain Perception." *The Rationalists: Descartes, Spinoza, Leibniz*. P. 257.

decir que ningún objeto material o ‘mecánico’ podría nunca explicar la mentalidad en su aspecto fenomenal, en la medida en que no tiene conocimiento intrínseco de lo material o lo ‘mecánico’. Es decir, el argumento completo de Leibniz está basado en la supuesta diferencia entre lo que se ve al interior del molino, que representa lo material, y las cualidades conocidas de lo mental. Pero esa diferencia, si Russell está en lo correcto, está basada en una confusión respecto a lo que sabemos de la materia. En realidad no sabemos lo suficiente como para alcanzar esas conclusiones.

I. La Teoría Russelliana Causal de la Percepción

A partir de *The Analysis of Mind* y en libros como *The Analysis of Matter*, *An Outline of Philosophy* y *Human Knowledge* Russell acepta y desarrolla una teoría causal de la percepción. El argumento para sostener que hay líneas causales que desembocan en la percepción emerge, en esencia, de la física y presupone el testimonio de otras mentes (Russell, 1927b, p. 98-99).

Podemos imaginar una serie de individuos, cada uno a una distancia pequeña de otro, de manera que todos estén percibiendo algún evento, digamos, un auto que se acerca. El preguntar a cada uno de ellos qué percibe mostrará que hay ligeras diferencias en sus percepciones, todas las cuales se vuelven explicables una vez que se aceptan algunas hipótesis. Por ejemplo, que la luz siempre viaja la distancia más corta entre dos puntos, que la misma se origina desde un ‘centro’, etc. Todo ello nos inclina a aceptar la existencia de varias líneas causales separables, todas las cuales tienen un origen en común y desembocan en el aparato perceptual de los varios individuos.

La importancia de las perspectivas para establecer la hipótesis plausible de ‘centros’ desde donde se irradian diversas líneas causales separables lleva a Russell a insistir que en lugar de individuos podríamos colocar cámaras o grabadoras, las cuales tomarán fotografías o grabarán sonidos que pueden contrastarse con cámaras y grabadoras a diversas distancias, Esto mostrará como los diferentes datos que son captados en diversas perspectivas son mejor explicados como productos de una línea causal, cuyo eslabón final es la cámara o grabadora, y cuyo origen (relativo al evento que se deseaba registrar) es algún cúmulo de eventos los cuales ‘irradian’ desde un centro. Todo esto, obviamente es consistente y explicable por la física, la cual Russell asume como verdadera en sus lineamientos más generales.

Supongamos por un momento que el realismo ingenuo, en su versión más radical, es el caso. Bajo esta postura tendríamos contacto inmediato, no mediado, con los objetos de percepción. Pero si esto es así varias cosas se vuelven increíbles, sino es que incoherentes. Para empezar sería increíble que los objetos que percibamos, digamos visualmente, sean los que están directamente frente a nuestros ojos ¿por qué esos objetos y no otros dado que la información se transmite automáticamente? Además, sería increíble que, por ejemplo, si un objeto está frente a mis ojos, perciba ese objeto y no cualquier otro que éste cubra. Por ejemplo, ahora mi laptop está frente a mis ojos y no percibo mi taza detrás de mi laptop. ¿Pero por qué tendría que percibir los objetos que están más cercanos a mis ojos que otros que están más distantes? Desde luego no es lógicamente imposible que así ocurriera pero si se vuelve a priori sumamente improbable, mientras que desde una teoría causal que asume la verdad de la física esto es precisamente lo que esperaríamos. Y ¿qué hay de fenómenos como reflejos de luz o neblina o cuando me pongo gafas oscuras?

Por otro lado los enigmas clásicos de las alucinaciones permanecen inexplicables bajo esta postura, mientras que en la medida en que introducimos la complejidad causal de cadenas de eventos que desembocan en alguna región espacio-temporal esos fenómenos se vuelven explicables.

Parece entonces que si asumimos el realismo ingenuo en su versión más radical, como la transmisión inmediata de información perceptual, entonces no podemos explicarnos toda clase de cosas sobre por qué percibimos como lo hacemos. No sé si el realismo ingenuo desemboca de manera directa en una contradicción, constituyendo un *reductio ad absurdum*, pero como mínimo la adopción de esa teoría es completamente inservible para explicarnos por qué percibimos como lo hacemos.

La teoría causal de la percepción tiene poder explicativo y presenta una visión coherente del mundo. Explica satisfactoriamente también porque las inferencias que van de las sensaciones a los objetos que suponemos las causan son falibles, en la medida en que algún proceso análogo en nuestros propios cuerpos puede ser suficiente para estimular algún elemento del sistema nervioso central produciendo tales sensaciones.

Evidentemente hay teorías que caen bajo el rubro de realismo directo que son más sofisticadas y que emplean la noción de intencionalidad para dar cuenta de la percepción, sin embargo entrar en disputa con esas posturas nos llevaría demasiado lejos de nuestro objetivo. Si bien podemos al menos observar que no es claro qué razones tendríamos para motivar posiciones así si podemos servirnos de una teoría que es no sólo consistente con la física, sino armoniosa con la misma y lo único que pide es suponer que nuestra percepción de objetos externos a nuestra mente sea indirecta. Un posible problema aquí podría ser la

suposición de que una visión así del mundo tiene que ser incoherente, pese a tener una distinguida tradición en la historia de la filosofía. Sin embargo, como veremos en la sección del problema del lindero, el cargo de incoherencia es injustificado.

II. La Lógica Matemática de Russell y La Estructura del Mundo Físico

Es necesario tener en mente algunas definiciones de forma que podamos caracterizar lo que son una estructura concreta y una estructura abstracta. Para hacer eso voy a repasar primero algunos temas Russellianos estándar en filosofía de las matemáticas y armado con eso voy a caracterizar las nociones relevantes. Vamos a seguir a Russell en su *Introduction to Mathematical Philosophy*, chp. IV y a Ioannis Votsis (2003).

Logicismo Russelliano: Aspectos Relevantes⁶⁶

La filosofía de las matemáticas de Russell constituye una temática amplia y técnica que abarca cuestiones tan variadas como los son la naturaleza de las matemáticas, problemas de ontología referentes a los objetos matemáticos y cuestiones de epistemología de las matemáticas. Por razones de espacio, aquí nos limitaremos a un esbozo informal.

El logicismo puede caracterizarse mediante las siguientes tesis⁶⁷:

⁶⁶ Una versión anterior de esta discusión aparece bajo mi autoría en mi blog (con Carlos Romero) de difusión de la filosofía La Pluralidad de los Mundos www.pluralidad.tk; me pareció importante introducir estos elementos para poder motivar la discusión en filosofía de la física-mente de corte Russelliano, que involucra extensamente su lógica matemática.

⁶⁷ Voy a seguir en esto una variante de la formulación de Mark Sainsbury. Para un estudio técnico y adecuado del tema véase la sección *Mathematics* en su libro *Russell* así como *Russell's Hidden Substitutional Theory* de Gregory Landini.

- i) Toda verdad matemática puede expresarse en un lenguaje cuya totalidad de expresiones son estrictamente lógicas; en otras palabras, toda verdad matemática puede expresarse como una verdad lógica.
- ii) Toda verdad matemática *es* una verdad lógica.
- iii) Las proposiciones de las matemáticas puras/lógica pura son completamente generales.
- iv) Toda verdad matemática, una vez expresada como proposición lógica, es deducible de un número pequeño de axiomas y reglas lógicas.⁶⁸

De manera informal, podemos decir que el logicismo es la tesis de que la lógica y las matemáticas son idénticas, al menos en la versión de Russell. En *Los Principios de las Matemáticas*, Russell ofrece una definición ya clásica de las matemáticas como sigue: las matemáticas son todas aquellas proposiciones de la forma ‘si p entonces q ’, donde p y q no contienen constantes, excepto constantes lógicas y las mismas variables aparecen en p y en q . Por lo mismo podemos sostener que Russell caracteriza las proposiciones de las matemáticas como hipotéticas, y especifica que deben estar conformadas por vocabulario

⁶⁸ La verdad es que es posible tener una versión más débil de logicismo, esto es, una versión en la que no se insista en que todos los enunciados de las matemáticas son consecuencias deductivas de axiomas lógicos en un sistema recursivamente axiomatizable, etc. Existe la idea de que el logicismo fracasó en virtud del teorema de incompleción de Gödel, cuyo contenido, de manera informal, es que en cualquier teoría consistente, recursivamente axiomatizable suficientemente poderosa para generar la aritmética de Peano existirá al menos una proposición verdadera pero no demostrable a partir de los axiomas del sistema. Pero esto último era más una esperanza de los primeros logicistas, Frege y Russell, que la esencia del proyecto. La esencia del proyecto, para Russell, no es más que la identidad entre las proposiciones de la lógica de orden superior y las matemáticas. Que esas sean completamente generales: siendo las únicas constantes las constantes lógicas y con variables para individuos y propiedades que no hagan referencia a dominios específicos. La axiomatización de la matemática es excelente evidencia de la verdad del logicismo; si ésta no puede completarse, no deja su gran éxito de ser evidencia si bien menos definitiva. Sobre los prospectos del logicismo, Fregeano y Russelliano, estos días véase Klement (2012), Landini (1998), Rayo (2002), Wright y Hale (2001).

estrictamente lógico. Sería quizá interesante diferenciar entre la tesis de los hipotéticos y el logicismo, por cuanto muchas teorías podrían caracterizarse como teniendo proposiciones de la forma ‘ p entonces q ’. Empero, lo particular de la teoría de Russell es que no sólo deben las proposiciones matemáticas ser de esa forma, sino que también deben poder expresarse en un vocabulario estrictamente lógico: completamente general, salvo las constantes lógicas y ser (en principio) capaces de ser deducidas lógicamente en forma válida desde un sistema recursivamente axiomatizable.

Algo que motiva a Russell a sostener que el logicismo es la concepción correcta de las matemáticas es la plausibilidad de varias reducciones de grandes partes de las matemáticas a cuerpos axiomáticos. Un cuerpo axiomático es un conjunto de enunciados que afirman que los objetos de la teoría en cuestión tienen ciertas propiedades o mantienen entre si ciertas relaciones. Así, por ejemplo, las geometrías pueden caracterizarse completamente mediante sus axiomas. Igualmente es posible reducir grandes partes de la matemática pura tradicional a la teoría de los números naturales⁶⁹ y estos a los axiomas de Peano. Los axiomas de Peano son los siguientes:

- i) 0 es un número.
- ii) El sucesor de todo número es un número.
- iii) Ningunos dos números tienen el mismo sucesor.
- iv) 0 no es el sucesor de ningún número.
- v) Cualquier propiedad que pertenezca al 0 y al sucesor de cualquier número que tenga la propiedad, pertenece a todos los números: $\forall P ((P(0) \ \& \ \forall k (P(k) \rightarrow P(k + 1))) \rightarrow \forall n P(n))$. Este es el así llamado ‘principio de inducción matemática’.

⁶⁹ Russell, Bertrand. *Introduction to Mathematical Philosophy*. Dover. 1993.

Cualquier estructura abstracta que satisfaga esos axiomas puede ser considerada como similar a la serie de los números naturales salvo que los referentes de los naturales en esta última son clases de clases de elementos. Es el hecho de que tal serie, en toda su complejidad, pueda ser así caracterizada, sugiere que tal vez los axiomas mismos puedan reducirse a expresiones lógicamente más básicas. Hay tres términos primitivos en tal serie: cero, número y sucesor, y el logicista puede mostrar que existe una caracterización de los mismos en un lenguaje puramente lógico, de manera que rescata sus propiedades. Es muy importante señalar, antes de avanzar más, que la lógica, tal como Russell la entendía, no comprendía solamente la lógica de primer orden (cuantificación sobre individuos), sino también la lógica de orden superior y lo que él llamaba la teoría de clases. Hoy en día las colecciones de objetos se estudian formalmente desde la teoría de conjuntos. Los conjuntos y las clases Russellianas son parecidos salvo que las clases están asociadas con propiedades de una forma más natural. En el logicismo Russelliano las clases son ficciones lógicas que desaparecen en último análisis, por lo que no existen compromisos ontológicos con clases.

Los términos primitivos se pueden explicar cómo lo quiere el logicista si empleamos conjuntos y agregamos a la lógica de primer orden el signo ‘ \in ’ para indicar la relación de pertenencia entre conjuntos y sus elementos⁷⁰. Un conjunto se puede definir de

⁷⁰ Hay que aclarar que *Principia Mathematica* no tiene la relación de pertenencia como primitiva; *Principia* es una teoría de tipos ramificada en donde ciertas funciones proposicionales tienen clases asociadas y otras no, con principios que impiden generar clases que violen el principio del círculo vicioso. Para evitar que se puedan generar expresiones mal formadas por alguna especie de círculo vicioso, Russell introduce la estipulación de que, en el lenguaje formalizado de *Principia Mathematica*, exista una jerarquía de tipos y órdenes. El objetivo de la jerarquía es restringir la cuantificación de tal manera que afirmaciones sobre el todo de una colección, a menos que sea parte del todo del rango de argumentos significativo de una función proposicional, sea un sinsentido. De acuerdo con Russell un tipo es, entonces, el rango de significatividad de una función proposicional. Dentro de su rango están los argumentos para los cuales la función tiene valores

manera extensional, enumerando sus elementos, o de manera intensional, especificándolos mediante una propiedad que todos posean. Dos conjuntos pueden ser iguales o uno puede ser un subconjunto de otro. Un conjunto A es igual a un conjunto B si y sólo si $\forall x (x \in A \Leftrightarrow x \in B)$ y un conjunto A es subconjunto de B si y sólo si $\forall x (x \in A \rightarrow x \in B)$. Luego, podemos definir ‘número’ mediante el uso del producto cartesiano. El producto cartesiano A x B de dos conjuntos A y B es el conjunto de todos los pares ordenados cuyo primer componente es un miembro de A y su segundo componente un miembro de B. En símbolos $A \times B = \{(x, y) \mid x \in A \ \& \ y \in B\}$ Teniendo en cuenta al producto cartesiano podemos especificar una relación, llamada ‘función’. Una función F es un subconjunto de un producto cartesiano, en símbolos $F \subseteq A \times B$. Llamamos al conjunto A el dominio de F, y al conjunto B el contra-dominio de F. A cada miembro del dominio corresponde uno y uno sólo del contra-dominio⁷¹. En este caso específico nos interesa tener a la mano la definición de función biyectiva. Una función biyectiva, además de satisfacer las propiedades de toda función, cumple con que su rango⁷² es igual a su contra-dominio y con que $f(x) = f(y)$ si y sólo si $x = y$.

Decimos que un conjunto es similar a otro si y sólo si hay una función biyectiva entre ambos. Así, Russell define el número de una clase como la clase de todos los

y puede ser verdadera o falsa. Otra asignación de argumentos la haría un sinsentido. De esta manera se puede, en el lenguaje formal de *Principia Mathematica*, impedir la formación de expresiones que pequen en contra del principio del círculo vicioso. Esto efectivamente permite llevar a cabo a logicización de las matemáticas en *Principia Mathematica*. El problema es que se cuelan dificultades, dado que no es claro por qué la teoría de los tipos habría de ser considerada como parte de la ‘lógica’, tomando ‘lógica’ en el sentido amplio de Russell.

⁷¹ Esto se puede expresar sin la palabra ‘uno’, de tal que no hay circularidad, ni se asume la unidad en la definición de número : $\forall x(x \in A \rightarrow \exists y((y \in B \ \& \ Fyx) \ \& \ \forall z(z \in B \ \& \ Fzx \rightarrow z = y))$

⁷² El conjunto de valores que arroja la función evaluada en ‘x’, esto es, F(x).

conjuntos similares a un conjunto dado. Esta definición es conocida como la definición Frege-Russell de número cardinal⁷³, y fue descubierta por ambos de manera independiente. Russell la presentó por primera vez en su *The Logic of Relations* en 1901. Así, por ejemplo, el 2 quedaría definido como ‘la clase de todos los pares’, el 4 como ‘la clase de todos los cuádruples’, etc. Cualquier par, entonces, satisface la siguiente expresión lógica:

$$\exists x \exists y (x \neq y \ \& \ \forall z (Pz \Leftrightarrow (z = x \vee z = y)))$$

Tenemos entonces que el número cardinal (Russelliano) de una clase es el conjunto de todos los conjuntos similares al mismo. Así, la clase $\{\emptyset, \{\emptyset\}\}$ y la clase $\{\text{Russell}, \text{Wittgenstein}\}$ son similares en tanto existe una biyección entre ambas. Si tomamos todas las clases similares a esas lo que tenemos es otra clase, la clase de todos los objetos x , y diferentes y tales que cualquier otro objeto z es idéntico a x o a y . Esa clase es, según Russell, el número cardinal 2. Una clase tiene 2 elementos si y sólo si es ella misma elemento de 2 así construido.

Se puede observar que en ningún punto en las definiciones se introdujo ningún término numérico y se empleó únicamente el vocabulario de la lógica y la teoría de conjuntos. Igualmente es posible definir ‘cero’ empleando la definición de número y ‘sucesor’ empleando cuantificación y la definición de subconjunto. Parece que efectivamente es posible, en el logicismo de Russell, llevar el análisis más lejos que los axiomas de Peano. De hecho *Principia Mathematica* es suficientemente fuerte para

⁷³ Shapiro, 2007.

recuperar gran parte de las matemáticas clásicas⁷⁴ y sus problemas son más de corte filosófico.

Análogamente, una relación o estructura concreta no es más que una clase específica i.e. con los elementos dados de n -adas con un orden; mientras que una relación-abstracta o estructura-abstracta es una clase de clases de n -adas ordenadas similares: de estructuras o relaciones concretas similares. Sea un ‘campo’ la clase de aquellos elementos en el dominio y codominio de alguna relación. Dos relaciones (estructuras) concretas, P y Q , son similares si hay una biyección S cuyo dominio es el campo de P y codominio el campo de Q . Luego, se sigue particularmente que para toda x e y que pertenecen a P tal que xPy , entonces para toda w y z que pertenecen a Q , wQz . Podemos ver que si conocemos alguna propiedad estructural de alguna relación R y sabemos que la relación S es similar a R , entonces S tendrá la misma propiedad estructural. Esto da una idea de lo que queremos decir por estructura.

La idea intuitiva de mapa nos permite elaborar: Podemos generar un mapa de alguna relación y decir que dos relaciones tienen la misma estructura si satisfacen el mismo mapa i.e. si tomando a los términos de un relación y sustituyéndolos por los de la otra y dada las características de ambas relaciones ambas ‘interpretan’ el mismo mapa.

Algunos ejemplos de mismidad de estructura son clásicos, por ejemplo, el isomorfismo de tipo entre frases como ‘Sócrates ama a Platón’ y ‘Metternich odia a Napoleón’. O ejemplos famosos del *Tractatus Logico Philosophicus* como la relación entre la partitura, el disco de gramófono y la pieza musical.

⁷⁴ Véase Boolos, 1999.

Sobre la importancia de la estructura para la filosofía Russell señala:

Ha habido una enorme cantidad de especulación en la filosofía tradicional que pudo haberse evitado si la importancia de la estructura, y la dificultad de ponerse tras de ella, hubieran sido entendidas. Por ejemplo, comúnmente se dice que el espacio y tiempo son subjetivos, pero tienen contrapartes objetivas; o que los fenómenos son subjetivos, pero son causados por cosas en sí mismas, que deben tener diferencias *inter se* que corresponden a las diferencias en los fenómenos que ocasionan. Donde tales hipótesis se hacen, generalmente está el supuesto que podemos saber muy poco de las contrapartes objetivas. De hecho, sin embargo, si las hipótesis como señaladas son correctas, las contrapartes objetivas formarían un mundo con la misma estructura que el mundo fenomenal y permitiéndonos inferir del mundo fenomenal la verdad de todas aquellas proposiciones que pueden ser enunciadas en términos abstractos y se saben verdaderas de los fenómenos.(1919, p.61)⁷⁵

De nuevo en *Human Knowledge* Russell señala de forma explícita que supuestos son requeridos para que el análisis estructural nos de información del mundo físico:

Si aceptamos la máxima ‘misma causa, mismo efecto’ y su consecuencia ‘diferente efecto, diferente causa’. Si este principio es considerado como válido, podemos inferir de una sensación compleja o series de sensaciones, la estructura de su causa física pero nada más, excepto que las relaciones de vecindad [espacio-temporal] deben ser preservadas; i.e., causas vecinas deben tener efectos vecinos. (Russell, 1948, p. 254)⁷⁶

En virtud de que Russell acepta esas máximas, y lo hace desde tan temprano como *The Problems of Philosophy*, es fácil ver que se desprende de ello su postura respecto a la

⁷⁵ “There has been a great deal of speculation in traditional philosophy which might have been avoided if the importance of structure, and the difficulty of getting behind it, had been realised. For example, it is often said that space and time are subjective but they have objective counterparts; or that phenomena are subjective, but are caused by things in themselves, which must have differences *inter se* corresponding with the differences in the phenomena to which they give rise. Where such hypotheses are made, it is generally supposed that we can know very little about the objective counterparts. In actual fact, however, if the hypothesis as stated were correct, the objective counterparts would form a world having the same structure as the phenomenal world, and allowing us to infer from phenomena the truth of all propositions that can be stated in abstract terms and are known to be true of phenomena.” *Introduction to Mathematical Philosophy*, p. 61.

⁷⁶ “... if we accept the maxim “Same cause, same effect” and its consequence, “Different effects, different causes.” If this principle is regarded as valid, we can infer from a complex sensation or series of sensations the structure of its physical cause, but nothing more, except that relations of neighborhood must be preserved; i.e., neighboring causes have neighboring effects.” *Human Knowledge*, p. 254.

filosofía de la física y la filosofía de la ciencia en general, esto es, el realismo estructural epistémico. Bajo el supuesto de la verdad en líneas generales de la física y adoptando el testimonio, así como los datos recogidos por observadores meramente posibles desde ciertas perspectivas, se desprende una teoría causal de la percepción.

En la teoría causal de la percepción, hay eventos no percibidos y perceptos, los cuales son transparentes a la persona en la medida en que tiene conocimiento no inferencial de los mismos. Un percepto es similar a un *quale*, en tanto tiene contenido experiencial, cualitativo, etc. Sin embargo hay que remarcar que, dado que Russell rechaza al sujeto desde *The Analysis of Mind* (1921), el concepto de '*qualia*' como una cualidad de la experiencia del sujeto no describe de manera acertada lo que Russell tiene en mente. De manera que un percepto podría definirse como un evento que es conocido sin inferencia (en la filosofía tardía de Russell el análisis de 'conocer' también elimina al sujeto, pero no entraré en detalles para no desviarnos del tema).

En cualquier caso, dada esta teoría, se localiza al percepto, por asunciones causales naturales, en el cerebro de quien percibe. Pero dado que este percepto es el eslabón último, relativo a los intereses de quien percibe, de una cadena causal, entonces todo conocimiento del mundo físico, cuando éste consiste en todo aquellos eventos que no son perceptos, es inferencial. Sin embargo, bajo las máximas señaladas, puede saberse gran cosa sobre la estructura del mundo físico. Esto que se sabe es la física. La física, por lo discutido, caracteriza sus objetos en términos de relaciones extrínsecas. Inclusive propiedades como la masa son definidas extrínsecamente y de los objetos de la física sólo sabemos sus

propiedades en tanto miembros de una compleja estructura. Pero esto deja abierto un espacio para propiedades intrínsecas.

Esta postura tiene definitivas ventajas en el debate en filosofía de la ciencia y en el presente debate respecto al problema-mente cuerpo, es lo que provee a Russell de su intuición principal. En cualquier caso es necesario extraer un par de consecuencias más de la postura de Russell para poder evaluarla *in toto* y ver qué puede hacer para superar los obstáculos que se han delineado a lo largo de este artículo.

III. Espacio-Tiempo y Espacio Perspectival

En esta sección sólo voy a notar superficialmente otras características de la postura de Russell. La primera está relacionada a la localización de lo mental, del *qualia*. ¿Cómo puede el *qualia* tener locación espacial? Un argumento central, que aparece en (Russell, 1927a) y en (Lockwood, 1989) está relacionado al abandono de un espacio euclideo con una dimensión temporal separable y su sustitución por un espacio Minkowski, el cual no posee una dimensión temporal separable⁷⁷. Lo importante en este caso del espacio Minkowski es que sustituye espacio y tiempo por espacio-tiempo, y por ende sustituye el hablar sobre puntos e instantes por hablar de puntos-instantes. Bajo esta concepción, si un evento está en el tiempo entonces está en el espacio (y viceversa). Pero una mayoría de

⁷⁷ Por separabilidad entiéndase aquí el hecho de que la métrica de un espacio Minkowski i.e. la función de distancia entre puntos (eventos) que satisface desigualdad del triángulo, simetría y no-negatividad, va del producto cartesiano $E * E$ de eventos espacio-temporales a los (números) reales. Es decir, la métrica de esos espacios está definida sobre puntos-instantes al contrario de la métrica de espacios Newtonianos y Galileanos, que está definida ya sea sobre puntos, para hiperplanos de simultaneidad, o instantes, para la dimensión temporal. La distancia invariante desde cualquier marco de referencia en un espacio Minkowski se conoce como intervalo. Véase por ejemplo Balashov, 2010.

nosotros concederíamos que nuestros *quale* están en el tiempo. Algún *quale* de dolor, por ejemplo, ocurre en ciertos tiempos solamente. Dada la teoría causal de la percepción, y presuponiendo la física y la fisiología, es evidente que algún evento físico impacta alguna superficie corpórea y es registrado como una sensación. Pero entonces, por economía y por razones de continuidad causal, lo más razonable es situar a esa sensación, *quale*, percepto, etc., en el cerebro de quien percibe i.e. como miembro de una estructura concreta material que también empalma una estructura concreta mental-funcional (psicológica).

Es importante, a continuación, distinguir entre el espacio perspectival y el espacio físico. El espacio perspectival es simplemente el espacio fenoménico, aquel en donde se observan de manera directa y sin inferencia ciertas cualidades, como el campo visual de cualquiera. Un conjunto de *qualia* de un árbol, en el espacio perspectival, estará a cierta distancia del percepto, como cualidad en ese espacio perspectival, que tiene el sujeto de su propio cuerpo. Sin embargo ese mismo *quale*, en el espacio físico, se encuentra literalmente en la región de la cabeza del sujeto: de nuevo concediendo la teoría causal de la percepción. Es evidente que la posición del *quale* en el espacio perspectival está relacionada a la posición del objeto causante del *quale* en el espacio físico i.e. existe cuando menos una relación de *embedding*⁷⁸ entre las posiciones espaciales de los objetos en el espacio perspectival y las posiciones de los objetos materiales en el espacio físico.

Aceptando el realismo estructural epistémico, puede inferirse, a partir del espacio perspectival, que dado que en este se manifiestan ciertas distancias entre el *quale* del árbol

⁷⁸ Decimos de que se da un *embedding* entre dos conjuntos A y B si y solamente si existe una función inyectiva $f: A \rightarrow B$, tal que para toda x, y en la relación R en A; $f(x), f(y)$ están en la relación R' en B.

y quien percibe, un isomorfismo (o cuando menos *embedding*) estructural valdrá para el espacio físico. Si esto es así la razón por la cual podemos navegar tan efectivamente en el espacio físico empleando el espacio perspectival de nuestros *qualia* es porque inconscientemente hemos aprendido a traducir propiedades de un espacio a otro, en la medida en que nuestras estimaciones de distancias, nociones de dimensión, etc., tienen las mismas propiedades formales en ambos espacios.

IV. El Problema de Linderos

¿Es coherente suponer que nuestro conocimiento de la materia externa a nuestro cuerpo es indirecto al grado en que Russell y otros suponen? En esta sección voy a considerar un experimento mental esgrimido por Russell (1927^a) que pretende establecer la coherencia de esta visión del mundo.

Russell plantea su experimento mental en el capítulo *Electrons & Protons* de su *The Analysis of Matter*. Nos pide que imaginemos una esfera dentro de la cual hay emisiones de energía. Nuestra posición respecto al interior de la esfera es de ignorancia. Piensa él que en ese caso tendremos varias hipótesis respecto a lo que ocurre en el interior de la esfera. Siempre y cuando sean consistentes y predigan lo que va a ocurrir en la superficie de la misma podemos considerar a todas legítimas. Luego Russell nos pide que imaginemos que esa esfera aumenta de tamaño, hasta que cubre todo hasta la superficie de nuestros cuerpos. En ese caso lo único que podríamos saber con certeza sería lo que ocurre en la superficie de los mismos y respecto a lo que cause los cambios en la superficie tendríamos que valernos de inferencias. Ésta, piensa Russell, es nuestra posición respecto al mundo físico.

Russell, efectivamente, replantea el problema del escepticismo como un problema de linderos, los linderos del cuerpo de quien percibe y tiene acceso al mundo físico vía sus superficies sensibles. Creo que su experimento mental muestra, entre otras cosas, que la idea tradicional de que sólo puedo tener acceso directo a los contenidos de mi conciencia es perfectamente coherente y sin duda la posibilidad de caer en el escepticismo no es una refutación de esta postura. Es decir, aun si la teoría causal de la percepción posibilita el escepticismo, no parece que eso por sí mismo constituya una demostración de incoherencia de la postura.

V. El Argumento Bidimensional en Contra del Fisicalismo

Hasta ahora los argumentos que se han avanzado a favor del dualismo no han logrado establecer gran cosa. Quizá los más respetables sean el argumento de la concebibilidad y el de Jackson, sin embargo no poseen suficiente fuerza como para forzarnos a abandonar supuestos preciados, como el del cierre causal del mundo físico. Sin embargo voy a explorar otro argumento bastante más sofisticado y que, considero, es en buena medida exitoso si se acepta cierta semántica. La consecuencia de su correctud y validez sería el fortalecimiento del caso en favor del monismo Russelliano. Este argumento ha sido propuesto por David Chalmers a lo largo de varias publicaciones. Aquí voy a basarme en su artículo *Consciousness & Its Place in Nature*.

Lo primero es notar que el tipo de materialismo (fisicalismo) que confronta este argumento es el llamado materialismo tipo-B por el propio Chalmers. En esta postura el materialista acepta que hay un hueco explicativo de lo material a lo mental. No es claro cómo es que lo material genera experiencias, etc. En virtud de ello este materialista acepta

que la proposición (P & - Q) donde P es la verdad física completa del mundo y Q alguna proposición fenomenal, es concebible, pero rechaza que sea metafísicamente posible. Es decir, considera que este es un mundo donde de hecho (P → Q) es un *necesario a posteriori*, al estilo de H₂O = Agua. Me parece que es difícil ser un materialista de otro tipo. Es decir, no encuentro mucho sentido a que se señale que (P & - Q) ni siquiera es concebible. Pero no exploraré esa posible objeción a otras formas de materialismo.

Regresando al argumento. La idea es que si (P & - Q) es concebible, debe haber algo en el mundo que la haga así. En este caso, podemos considerar nuestra situación epistémica antes de saber que H₂O = Agua. Parece que en ese momento no sabíamos en cuál de los diversos mundos posibles estábamos, al descubrir que H₂O = Agua, se reduce el rango de mundos posibles en los que sospechamos que estamos. Pero la ilusión de que otra cosa pudo haber sido agua ¿de dónde surge? No parece metafísicamente imposible que la sustancia XYZ hubiese dado lugar a aquello que tiene las propiedades del agua. Pero si 'H₂O = Agua' es verdad, entonces no es posible que 'XYZ = Agua'. Sin embargo pudimos haber estado en el mundo posible donde 'XYZ = Agua', es decir, donde ese mundo es @. Bajo esa situación, entonces ese mundo 'verifica' la verdad de 'XYZ = Agua'.

En este caso, nos dice Chalmers (Ibid, p. 255):

Pensamos en el mundo W como epistémicamente posible, como una manera en que el mundo podría ser. Cuando hacemos esto consideramos a W como actual. Cuando lo consideramos así puede que haga a alguna oración S verdadera o falsa. Si pensamos en el mundo XYZ como actual hace verdadero 'Agua no es H₂O'.⁷⁹

⁷⁹ "We think of W as an epistemic possibility: as a way the world might actually be. When we do this we consider W as actual. When we think of W as actual, it may make a given sentence S true or false. For example when thinking of the XYZ-world as actual, it makes "Water is not H₂O" true." *Philosophy of Mind: Classical and Contemporary Readings*, p. 255.

Podemos también considerar al mundo como contrafáctico, si ya hemos aceptado que se ha descubierto algún necesario *a posteriori*, entonces un mundo contrafáctico es aquel en donde XYZ hubiera sido agua. En esos mundos la proposición ‘H₂O no es agua’ es falsa. Y XYZ es simplemente una sustancia acuosa, pero no agua.

En una semántica bidimensional decimos:

- i) W considerado como @ hace verdadera a S ↔ W verifica a S.
- ii) W considerado como contrafáctico hace verdadera a S ↔ W satisface a S.

Luego asociamos a S con funciones que van de mundos a valores de verdad (intensiones).

- i) La intensión primaria de W es una función que es verdadera en W ↔ W verifica S.
- ii) La intensión secundaria de W es una función que es verdadera en W ↔ W satisface S.

Bajo el ejemplo en que ‘H₂O = Agua’ es S, S es falsa en su intensión primaria en el mundo XYZ, pero verdadera en su intensión secundaria.

Igualmente si S es ‘Sal ≠ NaCl’, entonces S es verdadera en su intensión primaria en el mundo donde ABC es la estructura química de la sal. Y S es falsa en su intensión secundaria, dado que el mundo ABC no satisface S.

Chalmers sugiere lo siguiente:

Cuando un enunciado S es concebible –esto es, cuando su verdad no puede ser descartada a priori- entonces hay algún mundo que verifica S.⁸⁰

⁸⁰ “When a statement S is conceivable –that is, when its truth cannot be ruled out a priori- then there is some world that verifies S.” *Philosophy of Mind: Classical and Contemporary Readings*, p. 256.

Ahora, sea P la verdad física completa, Q alguna verdad fenomenal, y bajo el criterio de concebibilidad de arriba tenemos:

- i) (P & - Q) es concebible.
 - ii) Si (P & - Q) es concebible, entonces un mundo verifica (P & - Q).
 - iii) Si un mundo verifica (P & - Q), entonces un mundo satisface (P & - Q) o la teoría russelliana es verdadera.
 - iv) Si un mundo satisface (P & - Q), el materialismo es falso.
-
- v) El materialismo es falso o la teoría russelliana es el caso.

¿Cómo se justifican las premisas? La premisa i) la concede el rival, el materialista que acepta el vacío epistémico. La premisa ii) se sigue de la aceptación del slogan de Chalmers, el cual parece plausible dado lo discutido. La premisa iii) como nota Chalmers, no es en general verdadera, por ejemplo en el caso del agua o la sal no se podía pasar de W verifica S a W satisface S. La idea es que Q, al ser una verdad fenomenal, si hemos aprendido la lección de Kripke que el modo de presentación de lo fenomenal es simplemente lo fenomenal: entonces si W verifica Q, entonces hay conciencia en W y luego W satisface Q. De nuevo, esto es porque el modo de presentación en este caso no es otra cosa que la conciencia y no puede ser, como en el caso de ‘Calor = Movimiento molecular’, que alguien esté en la ignorancia respecto a si hay conciencia como lo estaría respecto a si la sensación de calor es movimiento molecular u otra cosa. Luego, respecto a P. Si W verifica P entonces al menos la física de W debe ser estructuralmente idéntica a la

del mundo actual. Pero entonces la única forma en que no satisfaga S es que deje de fuera las propiedades intrínsecas de lo físico, esto es, la postura de Russell de que la física deja un espacio para las propiedades intrínsecas. Pero entonces, si P es el culpable de la falta de conciencia en W, pero lo único que P deja fuera son las propiedades intrínsecas entonces, como Russell pensaba, las propiedades intrínsecas de lo físico es lo fenomenal. La premisa iv) podría argüirse que es crucial al materialismo. La conclusión se sigue de las cuatro premisas de manera directa.

Nuestra discusión del dualismo en la primera parte de esta tesis nos da buenas razones para rechazar el dualismo como una solución viable al problema mente-cuerpo aun en sus aspectos mental-experienciales. La razón es que a grandes rasgos tenemos bien integradas diversas explicaciones del mundo físico; el mundo químico; el mundo biológico e inclusive el mundo psicológico en sus aspectos mental-funcionales. Si bien un reduccionismo duro respecto a ese tipo de dominios al dominio y leyes de la física es imposible en la práctica no hay nada que impida en principio suponer que las entidades y leyes de alto nivel están fundamentadas en las entidades y leyes de la física. Si esto es así, entonces la súbita postulación de un dualismo a nivel de seres de cierta complejidad psicológica cuyo único rol es explicar o intentar explicar los *qualia* es bastante insatisfactoria. Sin embargo el monismo Russelliano, que no postula ningún tipo de entidad adicional, es perfectamente compatible con esta visión de un mundo estructurado en términos de fundamentación y con que el nivel más fundamental sea el de la física, sin por ello comprometerse a que las entidades *sine qua non* de la experiencia sean ajenas a la física: ciencia de estructuras espacio-temporales-vectoriales en principio capaces de ser instanciadas por muchas cualidades intrínsecas diferentes.

Podemos entonces sintetizar el siguiente argumento a partir de la discusión previa:

- i) El materialismo es falso o el monismo-Russelliano es verdadero.
- ii) O el materialismo es verdadero o el dualismo es verdadero.
- iii) Pero el dualismo es falso. (Por las razones arriba dadas)

∴ Por lo tanto, el materialismo es verdadero y por (iv) el monismo Russelliano es el caso.

VI. En Síntesis

Lo crucial entonces, respecto a la postura de Russell, son las razones que nos da para rechazar la idea de que conocemos bien lo que es la materia. Si no conocemos bien lo que es la materia y la física caracteriza sus objetos mediante propiedades extrínsecas entonces no hay ningún obstáculo para identificar a los perceptos, sensaciones, *qualia*, con lo que está ocurriendo a nivel físico en mi cerebro y hay razones de sobra para hacerlo, como la teoría causal de la percepción y el hecho de que mis *qualia* deben estar en el espacio. Así, podemos decir que de una región del universo conocemos tanto las propiedades intrínsecas de la materia como sus propiedades matemáticas. De paso hay que aclarar que la postura de Russell no es puro estructuralismo, pues como él mismo lo señala en su autobiografía, siempre asumió copuntalidad espacio-temporal entre perceptos y no-perceptos⁸¹ (Russell, 1967, p. 413).

⁸¹ Esto permite superar la famosa objeción de Max Newman (1928) de que si lo único que conocemos del mundo físico es su estructura abstracta eso es equivalente a una afirmación sobre su cardinalidad i.e. el número de objetos en el mundo solamente. A cambio de abandonar un estructuralismo absoluto por uno

Hay, entonces, en esta teoría, una región del universo físico que conocemos de manera robusta: nuestra propia mente o cerebro. De nuevo hay que enfatizar que los ‘cerebros’ en esta concepción, no son la masa rosa, viscosa típicamente asociada a los mismos. Dado que esto último es sólo un percepto que en el espacio físico está localizado en la región de la cabeza de aquellos que miran a un cerebro. De los cerebros de otros tenemos sólo conocimiento inferencial, como aquello que, en la cadena causal que termina con nosotros, original ciertos *quale* de apariencia viscosa, y gris, etc. Pero en realidad el propio cerebro, en la postura de Russell, es algo conocido directamente y sin inferencia. Son simplemente los *token* de *qualia* (o perceptos) accesibles para un individuo a lo largo de su vida.

Hay que tener en cuenta que esta postura no es un mero dualismo de propiedades. Hay un dualismo de propiedades, si, pero este se da de manera insospechada. Lo que conocíamos de la materia era mero esqueleto y los *qualia* son la propiedad intrínseca de la materia. De esto obviamente no se sigue que toda región del universo tenga mentalidad. Curiosamente y no fue mencionado, para Russell lo crucial en lo mental es su aspecto funcional, dado que éste se conecta claramente con el desear, crear, volicionar, etc. En la medida en que sólo estructuras complejas son capaces de lo mental en su aspecto funcional y no sabemos a qué grado lo mental-experiencial permee el mundo físico, parece una elección razonable de Russell el suponer que no es el aspecto cualitativo sino el funcional

parcial, en dónde si bien del mundo físico conocemos en gran medida su estructura abstracta, también conocemos que ciertas relaciones espacio-temporales son las mismas a nivel perceptual y no-perceptual. Lo último permite individuar una clase de estructuras abstractas como aquellas que son estudiadas en física i.e. aquellas cuyas relaciones espacio-temporales son co-puntuales con perceptos ahí donde hay personas. Dado que sólo algunas estructuras físicas cumplen eso entonces podemos decir que conocemos del mundo más que su cardinalidad mínima.

lo que es característico de lo mental (Russell, 1956). Puesto que ignoramos, pero hay razones para sospechar, que el *qualia*, si en una región del espacio-tiempo es lo que rellena el esqueleto causal de lo físico, también lo hace en otras regiones.

VII. Diagnósticos

i) El argumento bidimensional

El argumento bidimensional es prometedor. Tiene, hasta donde veo, sólo dos dificultades. La primera es la premisa ii), que si bien suena plausible parece suponer demasiado sobre nuestra habilidad de olfatear siquiera lo epistémicamente posible. Respecto a la premisa iv) me parece demasiado fuerte. No estoy seguro de que esa premisa sólo le cause problemas a aquellos que defiendan el materialismo como una tesis necesaria. Alguien como David Lewis simplemente podría rechazar iv) sin perder el sueño. Finalmente, si concedemos las premisas y la opción está entre el dualismo o la teoría russelliana, la teoría russelliana me parece superior, en la medida en que es compatible con el cierre causal de la física, la influencia de lo mental, la indubitabilidad de los *qualia* y además no postula entidades misteriosas cuya interacción con lo físico resulte necesariamente incomprendible.

ii) Supuestos Pre-teóricos

Veamos qué nos dice la teoría de Russell respecto a nuestros supuestos intuitivos. En breve podemos decir que Russell no requiere abandonar ninguno y acepta IQ, IM, CCMF y UC. Su postura al ser anti-reduccionista no lo mete en conflictos con IQ. Dado que lo mental-experiencial es sólo la propiedad intrínseca de la materia entonces no se requiere un dualismo y la física sigue siendo causalmente cerrada. Russell acepta IM en la medida en que las voliciones, al ser las propiedades intrínsecas de eventos físicos, tienen efectos en el

mundo. No es un caso de sobre-determinación: hay un sólo evento, digamos mover el brazo y la volición de moverlo (cuando se cumplen ciertas condiciones antecedentes). Es un solo evento, sólo que el aspecto experiencial es la propiedad intrínseca de ese evento, mientras que las interacciones causal-funcionales son lo que corresponde a lo extrínseco, la estructura. Respecto a UC, hay un problema respecto a cómo reconciliarla con la multiplicidad de eventos físicos cuyas supuestas características intrínsecas dan lugar a la conciencia. Sin embargo, sin entrar en esa discusión, una posible salida es notar que los eventos físicos que componen el cerebro, como partículas, son hoy en día concebidas, por ejemplo en Teoría Cuántica de Campos, como campos y un campo tendría las propiedades requeridas de la unidad de la conciencia i.e. cumple que para cualquier punto en el campo existe un valor asociado. En el caso de campos cualitativos como el visual, eso es al menos compatible con la continuidad de cualidades en los mismos, por ejemplo.

iii) Los argumentos anti-fisicalistas

Hay que evaluar qué tan bien parada sale la posición Russelliana frente a los tres argumentos en favor del dualismo.

Respecto al argumento del conocimiento, lo único que hay que decir es que, contrario a la hipótesis, María no conocía todos los hechos físicos. Conocía sólo los hechos estructurales-disposicionales de la física. El problema es que el argumento fue diseñado bajo la ilusión, de acuerdo a esta teoría, de que sabemos bien lo que es la materia y es por eso que Jackson asume de entrada que el conocer alguna cualidad fenomenal es conocer algo no-físico.

El argumento del zombie en sus diferentes paráfrasis se responde como sigue: Efectivamente Chalmers tiene razón respecto a que, si caracterizamos a lo físico de cierta

manera, entonces los zombies parecen lógicamente posibles. Igualmente, bajo mi segunda paráfrasis, si F son todos los hechos físicos y L de todas las leyes de la física, caracterizadas bajo la concepción pre-russelliana, efectivamente F y L juntas no implican M. Respecto a la segunda paráfrasis, la dificultad es la misma: el aceptar cierta caracterización de lo físico que hace coherente el pensar con independencia lo mental-funcional y lo mental-experiencial. Como le señala Chalmers a Searle en un bien conocido debate⁸², el russelliano o quien acepta alguna variante de panpsiquismo, debe introducir la premisa adicional de que lo mental es lo que caracteriza las propiedades intrínsecas de lo físico. Una vez que se acepta eso se ve que los zombies no son lógicamente posibles y que F & L, revisadas russellianamente, implican M.

La respuesta russelliana al argumento de Kripke es más elaborada y no hay un acuerdo entre los estudiosos de su filosofía acerca de cómo iría más allá de algunos planteamientos generales. Un buen artículo es (Maxwell, 1978). Un esbozo de una plausible respuesta es como sigue, pero voy a dar varios. Notar que efectivamente es necesario que dolor sea idéntico a algún estado cerebral, en la medida en que el estado cerebral simplemente es el dolor. Pero no hay problema, en la medida en que no se desea insistir en que el dolor es sólo un modo de presentación que nos permite referir a un estado cerebral, como en las teorías eliminativistas de la identidad. Es decir, Russell no diría que algún estado cerebral de ‘dolor’ pudo haberse sentido diferente, llegando al absurdo en el que caían los teóricos clásicos de la identidad; él diría que el dolor simplemente es el estado

⁸² *Consciousness and The Philosophers* en la *New York Review of Books*
<http://www.nybooks.com/articles/archives/1997/may/15/consciousness-and-the-philosophers-an-exchange/>

cerebral, un aspecto que conocemos de manera directa, mientras que el cerebro, en la concepción del fisiólogo, es sólo conocido en tanto sus efectos en nosotros y su estructura.

Esto no impide que haya seres fisiológicamente diferentes pero capaces de experimentar sensaciones. Así como hay gamas de sensaciones, y posibles modalidades sensoriales que nosotros no podamos conocer, como el experimentar el sonar de un murciélago, igualmente es posible que hay diferentes modalidades de dolor que experimenten otros seres cuya estructura “cerebral” sea diferente. De nuevo, las teorías russellianas no son incompatibles, y de hecho son complementarias, con una versión débil de funcionalismo no-reduccionista. Una que acepte que en algún ser, el rol que cierto *qualia* ocupa en cierta estructura causal-funcional, es relevante para determinar su estructura psicológica, así como sus respuestas, etc, sin rechazar que si alguna sensación es idéntica a cierto estado cerebral, necesariamente lo es.

Otra forma de responder es notando que, en la medida en que la física y la fisiología lidian con estructuras complejas; entonces un estado cerebral es un tipo de estructura⁸³ mientras que una cualidad fenoménica o proto-fenoménica es un individuo. Claramente son diferentes. Sin embargo, en la medida en que las cualidades-en-ciertas-relaciones instancian en el mundo actual esas estructuras, entonces son cualidades-en-relaciones en el mundo actual son lo que da lugar a esas estructuras físicas, fisiológicas y causal-funcionales. Pero

⁸³ En ideología contemporánea, la física y la fisiología se ocuparían de ciertos conjuntos con relaciones particulares sobre los mismos, mientras que las cualidades serían los elementos del dominio. Esto puede expresarse también en la ideología de *Principia Mathematica*.

en otros mundos otras cualidades, como las cualidades alienígenas de Lewis⁸⁴, podrían dar lugar a esas estructuras.

Siempre y cuando las estructuras coincidan a nivel de isomorfismo, esos mundos serán físicamente idénticos, pero de eso no se seguirá que sean cualitativamente idénticos. Esto particularmente nos inclina a aceptar que, pese a que las cualidades tienen aspectos intrínsecos, estos no necesariamente corresponden al rol causal-funcional que tienen en el mundo actual. Pero después de todo eso no es tan extraño. Hay personas que experimentan “dolor” con placer y personas que prueban el chocolate con indiferencia. Plausiblemente, las cualidades intrínsecas tampoco cambian ahí, pero hay presentes ciertas relaciones causal-funcionales que sí.

⁸⁴ Véase Lewis, 1986.

CONCLUSIÓN

En este trabajo de filosofía abordé uno de los problemas tradicionales más complicados. El problema de la experiencia y su relación con el mundo físico desde una óptica russelliana.

Particularmente intenté sostener que:

- i) El dualismo no es una buena teoría en diferentes dominios de la metafísica de lo mental y ello nos da razones para sospechar de la misma en el dominio de la experiencia.
- ii) El dualismo tampoco es una buena teoría en el dominio de la experiencia, si bien tampoco es decididamente incoherente.
- iii) Hay ciertos supuestos pre-teoréticos que toda teoría de la mente debe acomodar o explicar con fuerza porque deben abandonarse.
- iv) El fisicalismo tradicional no puede acomodar bien esos supuestos.
- v) El fisicalismo tradicional deja un hueco explicativo de lo material a la experiencia que es insalvable.
- vi) La posición Russelliana puede acomodar esos supuestos y salvar el hueco explicativo dada su óptica particular respecto al problema de nuestro acceso epistémico al mundo físico.

Si lo que he venido diciendo es correcto, entonces del mundo físico a lo sumo conocemos su estructura abstracta, pero si eso es así, entonces esa estructura podría

instanciarse de múltiples maneras salvo restricciones de cardinalidad⁸⁵. La física entonces resulta ser, como sostienen los realistas estructurales ónticos (cf. Ladyman, 2007) una ciencia de estructuras, pero contrariamente a lo que ellos sostienen hay algo más en el mundo: cualidades fenoménicas o proto-fenoménicas⁸⁶. Los aspectos mental-funcionales de este mundo, en la medida en que requieren cierta complejidad estructural, son entonces también posibles en cualquier mundo en dónde estructuras isomorfas a los mismos sean posibles; de esta manera vemos cómo es que el fisicalismo russelliano se concilia con su monismo neutral.

Las cualidades del mundo, sean las que sean, formarán diferentes complejos estructurales e.g. las posiciones de una partícula a través del espacio-tiempo o redes neuronales en disposiciones de creencia y, dependiendo del arreglo de esas estructuras de las que sean miembros, esas cualidades serán o mentales o físicas. En nuestro mundo sabemos, porque tenemos evidencia de primera persona, que esas cualidades tienen características intrínsecas que, en el caso de personas y plausiblemente animales, son fenoménicas y plausiblemente en otros casos proto-fenoménicas. Pero lo esencial a que esas cualidades sean materiales o mental-funcionales dependerá del tipo de estructuras de las que sean elementos sin que por ello pierdan sus cualidades intrínsecas⁸⁷; aun cuando el notar o ser consciente de esas cualidades requiera que sean elementos de algún complejo mental-funcional. Esto es consistente con la posición agnóstica de Russell, de acuerdo a la cual no sabemos si las cualidades que instancian las estructuras físicas son como las

⁸⁵ Véase Newman, 1928.

⁸⁶ Para más sobre este punto véase Chalmers, 2013.

⁸⁷ Russell, 1915.

cualidades de la experiencia o similares; pero sí podemos reconocer los aspectos mentalfuncionales en otros desde la óptica de la tercera persona⁸⁸.

⁸⁸ Russell, 1927b.

BIBLIOGRAFÍA

- Agustín, Vicente. *On the Causal Completeness of Physics*. International Studies in the Philosophy of Science. 20 (2):149 – 171. 2006.
- Anderson Anthony; Savage Wade. *Rereading Russell: Essays on Bertrand Russell's Metaphysics and Epistemology*. University of Minnesota Press. 1989.
- Balashov, Yuri. *Persistence and Space-Time*. Oxford University Press. 2010.
- Boolos, George. *Logic, Logic and Logic*. Harvard University Press. 1999.
- Conee, Earl & Sider, Ted. *Riddles of Existence*. Oxford University Press. 2007.
- *Consciousness and The Philosophers* en la *New York Review of Books*
<http://www.nybooks.com/articles/archives/1997/may/15/consciousness-and-the-philosophers-an-exchange/>
- Chalmers, David. *Consciousness & Its Place in Nature* reprinted in *Philosophy of Mind: Classical & Contemporary Readings*. Oxford University Press. 2002.
- Chalmers, David. *The Conscious Mind*. Oxford University Press. 1996.
- Chalmers, David. *Philosophy of Mind: Classical & Contemporary Readings*. Oxford University Press. 2002.
- Creath, Richard; Friedman, Michael eds. *The Cambridge Companion to Carnap*. Cambridge University Press. United States. 2007.
- Dawkins, Richard. *The Extended Phenotype*. Oxford University Press. 1982.
- Dawkins, Richard. *The Selfish Gene*. Oxford University Press. Great Britain. 1976.
- Dennett, Daniel. *Darwin's Dangerous Idea*. Simon and Schuster. New York. 1995.
- Hale, Bob; Wright, Crispin. *The Reason's Proper Study: Essays towards a Neo-Fregean Philosophy of Mathematics*. Oxford University Press. 2001.

- Hawthorne, John. *Metaphysical Essays*. Oxford University Press. New York. 2006.
- Jacob, Pierre, "Intentionality", *The Stanford Encyclopedia of Philosophy* (Fall 2010 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/fall2010/entries/intentionality/>.
- Jackson, Frank. *Epiphenomenal Qualia*. *The Philosophical Quarterly*, Vol. 32, No. 127. (Apr., 1982), pp. 127-136.
- Jackson, Frank. *What Mary Didn't Know*. *Journal of Philosophy* 83. 1986.
- Klement, Kevin C. *Neo-Logicism and Russell's Logicism*. *Russell: the Journal of Bertrand Russell Studies*: Vol. 32: Iss. 2, Article 3. 2012.
- Kagan, Shelly. Open Yale Courses. *Death*. Lectures 3, 4 & 5. <http://oyc.yale.edu/philosophy/phil-176>
- Kripke, Saul. *Identity and Necessity*. en *Identity and Individuation*, editado por M. K. Munitz. New York: New York University Press. 1971.
- Kripke, Saul. *Naming & Necessity*. Cambridge: Harvard University Press. 1980.
- Ladyman, James; Ross, Don. *Everything Must Go: Metaphysics Naturalized*. *Oxford University Press*. United Kingdom. 2007.
- Landini, Gregory. *Russell*. Routledge. New York. 2011.
- Landini, Gregory. *Russell's Hidden Substitutional Theory*. Oxford University Press. New York. 1998.
- Lewis, David. *On The Plurality of Worlds*. Blackwell. 1986.
- Lewis, David. *Philosophical Papers Volume I*. Oxford University Press. USA. 1983.

- Linsky, Bernard. *Russell's Metaphysical Logic*. CSLI Publications. Stanford, California. 1999.
- Locke, John. *Essay Concerning Human Understanding*. Collins. 1968. Originally published in 1690.
- Lockwood, Michael. *Mind, Brain & The Quantum*. Blackwell. 1989.
- Lockwood, Michael. *What was Russell's Neutral Monism?* In Peter A. French, Theodore E. Uehling and Howard K Wettstein. Eds. *Midwest Studies in Philosophy VI: The Foundations of Analytic Philosophy*. University of Minnesota Press. 1981.
- Macías Bustos, Moisés. *Wittgenstein y la teoría russelliana del juicio en Wittgenstein en Español* ed. por Silvia Rivera y Alejandro Tomasini Bassols. Serie Filosofía. Ediciones UNLa. Buenos Aires. 2009.
- M. H. A. Newman. *Mr. Russell's Causal Theory of Perception*. *Mind* 37. Pp. 137 – 148. 1928.
- Maxwell, Grover. *Rigid Designators and Mind-Brain Identity*. Reprinted in *Philosophy of Mind: Classical & Contemporary Readings*. Oxford University Press. 2002.
- Monk, Ray. *Bertrand Russell: The Spirit of Solitude*. The Free Press. New York. 1996.
- Nagel, Thomas. *What Is It Like to Be a Bat?*. *The Philosophical Review* **83** (4): 435–450. 1974.
- Parfit, Derek. *Reasons & Persons*. Oxford University Press. 1984.
- Passmore, John. *100 Años de Filosofía*. Alianza Universidad Editorial. Madrid. 1981.

- Perry, John. *Knowledge, Possibility & Consciousness*. MIT Press. 1999.
- Uzquiano, Gabriel; Rayo, Agustín eds. *Absolute Generality*. Oxford University Press. New York. 2006.
- Russell, Bertrand. *A Critical Exposition of the Philosophy of Leibniz*. George Allen and Unwin. Great Britain. 1900.
- Russell, Bertrand. *An Outline of Philosophy*. Norton. 1927b.
- Russell, Bertrand. *Introduction to Mathematical Philosophy*. London: George Allen & Unwin. 1919.
- Russell, Bertrand. *Human Knowledge: Its Scope & Limits*. Simon & Schuster. 1948.
- Russell, Bertrand. *Logic and Knowledge*. Routledge. New York. 2001.
- Russell, Bertrand. *My Philosophical Development*. Unwin Books. London. 1959.
- Russell, Bertrand. *On Denoting*. *Mind*, New Series, Vol. 14 No. 56. (Oct. 1905), pp. 479–493.
- Russell, Bertrand. *Portraits from Memory*. London: George Allen & Unwind. 1956.
- Russell, Bertrand. *The Problems of Philosophy*. Oxford University Press. USA. 1997.
- Russell, Bertrand. *The Analysis of Matter*. Kegan Paul. 1927a.
- Sainsbury, Mark. *The Arguments of the Philosophers: Russell*. Routledge and Kegan Paul. 1979.
- Shapiro, Stewart. *The Oxford Handbook of the Philosophy of Mathematics and Logic*. Oxford University Press. 2007.
- Smart, J. J. C., "The Mind/Brain Identity Theory", *The Stanford Encyclopedia of Philosophy* (Winter 2012 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/win2012/entries/mind-identity/>.

- *The Rationalists Descartes, Spinoza & Leibniz*. Anchorage. 1974.
- Sider, Theodore. *Fourdimensionalism: An Ontology of Persistence & Time*. Oxford University Press. 2001.
- Sider, Theodore. *Writing the Book of the World*. Oxford University Press. New York, 2011.
- Simons, Peter. *Parts: A Study in Ontology*. Oxford University Press. New York. 1987.
- Sklar, Lawrence. *Space, Time and Space-Time*. University of California Press. USA. 1977.
- Votsis, Ioannis. *Is Structure not Enough?* *Philosophy of Science*, vol.70 (5): 879-890. 2003.
- Wittgenstein, Ludwig. *Tractatus Logico Philosophicus*. Routledge. 2001.