



Ideas y Valores

ISSN: 0120-0062

revideva_fchbog@unal.edu.co

Universidad Nacional de Colombia

Colombia

Lewis, David

Nuevo trabajo para una teoría de los universales

Ideas y Valores, vol. LXIV, núm. 158, agosto, 2015, pp. 249-279

Universidad Nacional de Colombia

Bogotá, Colombia

Disponible en: <http://www.redalyc.org/articulo.oa?id=80941032012>

- Cómo citar el artículo
- Número completo
- Más información del artículo
- Página de la revista en redalyc.org

redalyc.org

Sistema de Información Científica

Red de Revistas Científicas de América Latina, el Caribe, España y Portugal

Proyecto académico sin fines de lucro, desarrollado bajo la iniciativa de acceso abierto

TRADUCCIÓN

<http://dx.doi.org/10.15446/ideasyvalores.v64n158.49713>

Nota del traductor

El texto que se presenta aquí es la segunda parte de la traducción de *New Theory of Universals* y se corresponde con las últimas secciones del artículo original. En el anterior número de *Ideas y Valores* (157 de abril de 2015) se publicó la primera parte, que corresponde a la introducción y las dos primeras secciones del original.

DIEGO MORALES

Pontificia Universidad Católica de Chile - Santiago de Chile - Chile
dhmoralet@uc.cl

NUEVO TRABAJO PARA UNA TEORÍA DE LOS UNIVERSALES

[Segunda Parte]

DAVID LEWIS

Duplicación, superveniencia y mundos divergentes

De aquí en adelante solo hablaré de mi necesidad alrededor de la distinción entre propiedades naturales y no-naturales, o más o menos naturales. Debe entenderse que el trabajo que tengo preparado para una teoría de las propiedades adecuadamente discriminatoria puede ser un nuevo trabajo para una teoría de los universales o, en cambio, un esfuerzo para los recursos de un nominalismo adecuado.

Comienzo con el problema de analizar la duplicación. Estamos familiarizados con casos de duplicación aproximada, *i.e.* cuando utilizamos máquinas de copiado. Entendemos que si estas máquinas fuesen más perfectas de lo que son, las copias que hacen serían duplicados perfectos del original. La copia y el original serían iguales en tamaño, forma y la composición química de las marcas de tinta y el papel, iguales en

temperatura y alineación magnética y carga electrostática, iguales incluso en la disposición exacta de sus electrones y quarks. Esos duplicados serían exactamente iguales, diríamos. Estarían perfectamente igualados, serían cualitativamente idénticos, serían indiscernibles.

Pero no tendrían exactamente las mismas propiedades, en el sentido que le doy a esta palabra. Como en el caso de dos cosas cualesquiera, un sinfín de divisiones de clases las dividiría. Intrínsecamente, dejando fuera sus relaciones con el resto del mundo, serían iguales. Pero ocuparían diferentes posiciones espacio-temporales; podrían tener diferentes dueños, ser examinadas por primera vez en diferentes siglos, y así sucesivamente.

Por lo tanto, si deseamos analizar la duplicación en términos de propiedades compartidas, parece que primero debemos distinguir las propiedades *intrínsecas* (o “internas”) de las *extrínsecas* (o “externas”). De este modo, podremos decir que dos cosas son duplicados si y solo si tienen precisamente las mismas propiedades intrínsecas, sin importar qué tanto sus propiedades extrínsecas puedan diferir. Pero nuestro nuevo problema de dividir las propiedades en intrínsecas y extrínsecas no es más fácil que nuestra cuestión original de analizar la duplicación. De hecho, ambos están unidos en un ajustado círculo de interdefinibilidad. La duplicación es un asunto de compartir propiedades intrínsecas; estas propiedades son aquellas que nunca difieren entre duplicados. La propiedad P es intrínseca si y solo si, para dos cosas duplicadas cualesquiera, no necesariamente del mismo mundo, ambas tienen P o ninguna la tiene. P es extrínseca si y solo si hay algún par de duplicados en donde uno tiene P y el otro no.¹

-
- 1 Dada la duplicación, podemos subdividir las propiedades extrínsecas, distinguiendo casos puros de aquellos que presentan mezclas de extrínsecas e intrínsecas. Dividase las cosas, de este y otros mundos, en clases de equivalencia bajo relaciones de duplicación. Una propiedad puede dividir una clase de equivalencia, puede incluirla o excluirla. Una propiedad P es extrínseca, como dijimos, si divide por lo menos algunas de las clases. Tenemos cuatro subcasos. (1) P divide cada clase; entonces podemos llamar a P *puramente extrínseca*. (2) P divide algunas clases, incluye algunas y excluye ninguna; luego P es la disyunción de una propiedad intrínseca y una propiedad puramente extrínseca. (3) P divide algunas, excluye algunas e incluye ninguna; entonces P es la conjunción de una propiedad intrínseca con una propiedad puramente extrínseca. (4) P divide algunas, incluye algunas y excluye algunas; luego P es la conjunción de una propiedad intrínseca con una propiedad impuramente extrínseca como aquellas consideradas en el segundo caso, o equivalentemente es la disyunción de una propiedad intrínseca y una propiedad impuramente extrínseca, como aquellas consideradas en el tercer caso. También podemos clasificar las relaciones en intrínsecas y extrínsecas, pero de dos maneras diferentes. Tómese una relación diádica, *i.e.* una clase de pares ordenados. Llámese *intrínseca a sus relata* a una relación si y solo si cuando a y a' son duplicados (o idénticos) y b y b' son duplicados (o idénticos), entonces ambos o ninguno de los pares

Si confiáramos en que nuestra teoría física fuera exacta y exhaustiva, podríamos pensar en definir la duplicación en los términos de aquella. Creemos que los duplicados deben ser iguales en la disposición de sus electrones y quarks –¿por qué no establecer esto como una definición?-. Pero dicha definición no es un “análisis”. Presupone la física de nuestro mundo actual; no obstante, aquella es contingente y conocida *a posteriori*. La definición no se aplica a la duplicación en mundos posibles en donde la física es diferente, o a la duplicación entre mundos que difieren en su física, ni tampoco logra capturar lo que aquellos que ignoran la física quieren decir cuando hablan –como lo hacen– de duplicación.

El curso apropiado, sugiero, es analizar la duplicación en términos de propiedades compartidas; pero sin comenzar con las propiedades intrínsecas, sino más bien con las propiedades naturales. Dos cosas son cualitativamente duplicados si tienen exactamente las mismas propiedades perfectamente naturales.²

La física es relevante porque aspira a entregarnos un inventario de las propiedades naturales –no un inventario completo, quizás, pero un inventario lo suficientemente completo como para dar cuenta de la duplicación entre las cosas actuales–. Si la física tiene éxito en esta tarea, entonces la duplicación dentro de nuestro mundo es equivalente a la igualdad de la descripción física. Pero las propiedades naturales son las que importan, no la teoría que nos dice cuáles son. Si el materialismo

$\langle a, b \rangle$ y $\langle a', b' \rangle$ se encuentran en la relación. Llámese *intrínseca a sus pares* a un relación si y solo si, cuando los pares $\langle a, b \rangle$ y $\langle a', b' \rangle$ son ellos mismos duplicados, entonces ambos o ninguno de ellos se encuentra en la relación. En el segundo caso, una exigencia más fuerte se impone sobre los pares. Por ejemplo, podrían fallar en ser pares duplicados porque la distancia entre a y b difiere de la distancia entre a' y b' , aun cuando a y a' sean duplicados y b y b' también lo sean. Según la terminología tradicional, las “relaciones internas” son intrínsecas a sus *relata*; “relaciones externas” son intrínsecas a sus pares, pero no a sus *relata*; y las relaciones extrínsecas incluso para sus pares, como la relación de pertenecer al mismo dueño, quedan fuera de la clasificación. Nuestra definición de propiedades intrínsecas, en términos de duplicación, se asemeja cercanamente a la definición de “propiedades diferenciales” dada por Michael Slote en “Some Thoughts on Goodman’s Riddle” (1967) y en *Reason and Scepticism* (1970). Pero donde yo cuantifico sobre *possibilia*, Slote aplica modalidad a cuantificadores ordinarios, presumiblemente actualistas. Eso hace una diferencia. Una propiedad extrínseca puede diferir entre duplicados, pero solo cuando los duplicados habitan en mundos diferentes; ahí Slote contaría la propiedad como diferencial. Un ejemplo es la propiedad de ser una esfera que habita un mundo en donde hay cerdos o un cubo que habita un mundo en donde no hay cerdos. Para una mayor discusión sobre el círculo que va desde la duplicación a la intrinsecalidad y viceversa, véase mi “Extrinsic Properties” (1983).

- 2 Igualmente $\langle a, b \rangle$ y $\langle a', b' \rangle$ son pares duplicados si y solo si a y a' tienen exactamente las mismas propiedades perfectamente naturales, así como también b y b' , y asimismo las relaciones perfectamente naturales entre a y b son exactamente las mismas que entre a' y b' .

fuese falso y la física un completo fracaso, como es el caso de algunos mundos deplorables, aún habría duplicación en virtud de propiedades naturales compartidas.

En mi análisis, todas las propiedades perfectamente naturales resultan ser intrínsecas. Eso parece correcto. Lo opuesto no es verdadero. Las propiedades intrínsecas pueden ser disyuntivas, misceláneas y no naturales, siempre y cuando no difieran entre duplicados. Las propiedades perfectamente naturales comprenden una base para las propiedades intrínsecas; pero los compuestos booleanos arbitrarios de estas, por muy poco naturales que sean, siguen siendo propiedades intrínsecas. Por consiguiente, si adoptamos el tipo de nominalismo adecuado, que traza una distinción primitiva entre propiedades naturales y propiedades no naturales, esto no es lo mismo que trazar una distinción primitiva entre propiedades intrínsecas y extrínsecas. La primera distinción da paso a la segunda, pero no *viceversa*.

Así mismo, si adoptamos el tipo de nominalismo adecuado que comienza con una apropiada relación de semejanza parcial, esto no es lo mismo que tomar la duplicación en sí misma como primitiva. Nuevamente, la primera da paso a la segunda, pero no *viceversa*.

En cambio, si rechazamos el nominalismo y consideramos que todas las propiedades perfectamente naturales son aquellas que corresponden a los universales (en el sentido de que los miembros de la propiedad son exactamente aquellas cosas que instancian el universal), entonces todas las propiedades que correspondan a los universales son intrínsecas. Así también son todos los compuestos booleanos –disyunciones, negaciones, etc.– de propiedades que corresponden a universales. Los universales en sí mismos son intrínsecos *ex officio*, por así decirlo.

Pero aquí debo confesar que la teoría de los universales, para la cual ofrezco un nuevo trabajo, no puede ser exactamente la teoría de Armstrong, pues debe rechazar los universales extrínsecos, mientras que este autor los admite, aunque no como irreducibles (*cf.* Armstrong II 78-79). Pienso que él se las arreglaría mejor sin ellos, dados sus propios objetivos. (1) Ellos subvierten la deseada conexión entre el hecho de compartir universales y hechos mooreanos de igualdad total o parcial de naturaleza. Reconocidamente, hay algo así como semejanza en aspectos extrínsecos: las cosas pueden ser parecidas en los papeles que cumplen *vis-à-vis* otras cosas o en los orígenes de los cuales provienen. Pero dichas semejanzas no son las que queremos expresar cuando decimos de dos cosas que son del mismo tipo o que tienen la misma naturaleza. (2) Ellos subvierten la deseada inmanencia de los universales: si algo instancia un universal extrínseco, no es un hecho solo sobre esa cosa. (3) Ellos no son necesarios para la teoría de las leyes de la naturaleza de Armstrong; cualquier supuesta ley que conecte universales extrínsecos

de cosas puede ser equivalentemente reemplazada por una ley que conecte estructuras intrínsecas de sistemas más grandes que contienen dichas cosas como partes.

Por lo tanto, estoy satisfecho con decir que si hay universales, entonces los duplicados intrínsecos son cosas que tienen exactamente los mismos universales. No es necesario decir “[...] exactamente los mismos universales *intrínsecos*”, pues no deberíamos creer en otro tipo.

No solo la duplicación es de interés por derecho propio; también es necesaria a la hora de hacerse cargo de otros temas en metafísica. Por lo tanto, dichos temas crean una necesidad derivada de propiedades naturales. En lo que sigue consideraré dos temas en donde creo necesario hablar de duplicación: superveniencia y mundos divergentes.

*

Primero, superveniencia. Una tesis de superveniencia es la negación de variación independiente. Dada una ontología de *possibilia*, podemos formular dichas tesis en términos de diferencias entre individuos o mundos posibles. Decir que tal-y-tal superviene en tal-y-cual es señalar que no puede haber diferencia con respecto a tal-y-tal sin diferencia en relación con tal-y-cual. La belleza de las estatuas superviene a su forma, tamaño y color, por ejemplo, si dos estatuas, en el mismo o en diferentes mundos, nunca difieren en belleza sin a la vez diferir en forma, tamaño y color.³

Una tesis de superveniencia es, en un sentido amplio, reduccionista. Pero es una forma elemental de reduccionismo, que no se encuentra sobrecargada de dudosas negaciones de existencia, aseveraciones de prioridad ontológica o sobre traducibilidad. Uno quisiera decir que en algún sentido la belleza de las estatuas no es algo añadido a la forma, tamaño y color que los espectadores aprecian, pero sin negar que haya algo así como la belleza, sin aseverar que la belleza existe solo de una manera menos fundamental y sin comenzar a parafrasear adscripciones de belleza en términos de forma, etc. Una tesis de superveniencia parece capturar lo que un reduccionista cauto quiere decir.

Incluso si los reduccionistas debiesen ser cautos y aspirar a la traducción, todavía es una buena idea atender a la cuestión de la superveniencia. Pues, en caso de que la superveniencia falle, entonces ningún esquema de traducción puede ser correcto y no necesitamos andar chisholmeando (*Chisholm-ing away*) * en busca de uno. Por el contrario, si la superveniencia tiene éxito, entonces algún esquema correcto debe

.....
3 Para discusiones generales sobre superveniencia, véase Kim (1978).

* *Nota del traductor*: “*To Chisholm-away*” es una expresión utilizada para designar la práctica de hacer pequeñas alteraciones, de forma repetida, en una definición o ejemplo

existir; la pregunta que queda es si acaso existe un esquema correcto que sea menos que infinitamente complejo. Si la belleza es superveniente a la forma, etc., lo peor que podría pasar es que una adscripción de belleza sea equivalente a una disyunción incontablemente infinita de descripciones máximamente específicas de forma, etc., las cuales pueden en sí mismas involucrar infinitas conjunciones.

Las tesis interesantes de superveniencia usualmente involucran la noción de duplicación cualitativa que recién consideramos. Por consiguiente, podemos preguntar qué superviene o no sobre el carácter cualitativo del mundo entero a lo largo de toda su historia. Supongamos que dos mundos posibles son duplicados cualitativos perfectos, ¿deben ellos entonces tener exactamente las mismas distribuciones de probabilidad objetiva, las mismas leyes de la naturaleza, los mismos contrafácticos y las mismas relaciones causales? ¿Deben sus habitantes tener las mismas propiedades modales *de re*? Si ese es el caso, tiene sentido proseguir con distintos proyectos, como un análisis frecuentista de la probabilidad, un análisis de regularidad de las leyes de la naturaleza, o un análisis de semejanza comparativa de contrafácticos causales y de la modalidad *de re*. Si no es así, dichos proyectos están condenados desde un principio, y no sería necesario atender a los detalles de dichos intentos. Pero no podemos plantear estas preguntas acerca de la superveniencia a menos que podamos hablar de mundos duplicados. Y para hacer eso, como he sugerido, necesitamos propiedades naturales. (Nótese que si los mundos posibles obedecen a un principio de identidad de indiscernibles cualitativos, entonces todas estas tesis de superveniencia valen inmediatamente. Si no hay dos mundos que sean duplicados, entonces *a fortiori* no hay dos duplicados que difieran en cuanto a sus probabilidades, leyes o cualquier otra cosa.)

También podríamos preguntar si acaso el carácter cualitativo superviene en cualquier cosa menor. Por ejemplo, podríamos preguntar si acaso el carácter cualitativo global superviene en el carácter cualitativo local. Digamos que dos mundos son duplicados *locales* si y solo si ellos son divisibles en partes más pequeñas correspondientes, de tal manera que (1) las partes correspondientes de los dos mundos son duplicados y (2) la correspondencia preserva las relaciones espacio-temporales. (El sentido exacto depende, por supuesto, de lo que entendamos por “pequeño”). Si dos mundos son duplicados locales, entonces ¿deben ser duplicados *simpliciter*? ¿O podrían diferir en aspectos que no previenen la duplicación local –e.g. en relaciones externas, distintas de las

.....
para alcanzar mayor claridad y rigor en la exposición. Esta expresión fue acuñada en honor a Roderick M. Chisholm (1916-1999).

relaciones espacio-temporales en sí mismas, entre cosas separadas–? Nuevamente debemos darle sentido a la duplicación –esta vez tanto en lo grande como en lo pequeño– para siquiera hacer la pregunta.⁴

*

A continuación, mundos divergentes. Diré que dos mundos posibles *divergen* si y solo si no son duplicados, pero tienen segmentos temporales iniciales que son duplicados. Es así como nuestro mundo y otro podrían ser perfectamente idénticos hasta el año 1945 y seguir caminos separados desde esa fecha en adelante.

Nótese que no es necesaria la identidad de los tiempos a través de los mundos. Nuestro mundo, a través de nuestro 1945, duplica un segmento inicial del otro mundo; el segmento de este termina con un año que efectivamente se asemeja a nuestro 1945, pero es parte del tiempo de otro mundo, no parte de nuestro tiempo. Así mismo, no es necesaria la separación del espacio y el tiempo que contraviene la relatividad –tenemos segmentos temporales iniciales, de este y otro mundo, si tenemos regiones espacio-temporales limitadas por superficies espacialoides que cortan el mundo en dos–.

Yo distingo la *divergencia* de los mundos de la *bifurcación* (*branching*) de estos mismos. En la bifurcación, en vez de segmentos duplicados, uno y el mismo segmento inicial es presuntamente compartido como una parte común por dos mundos que se superponen. La bifurcación es problemática en aspectos en que la divergencia no lo es. Primero, porque un habitante del segmento compartido no podría hablar unívocamente sobre *el* mundo en que vive. ¿Qué sucedería si es que él dice que habrá una batalla naval mañana, cuando quiere hablar del futuro de su propio mundo, y uno de los dos mundos en los que vive tiene una batalla naval al día siguiente y el otro no la tiene? Segundo, porque la superposición de los mundos interfiere con el principio más saliente de la demarcación de estos, *viz.*, que dos posibles individuos son parte del mismo mundo si y solo si están unidos por alguna cadena de relaciones externas, *i.e.* de relaciones espacio-temporales (no conozco otro ejemplo). Ninguna de estas dificultades parece insuperable, pero es mejor si ambas son evitadas. Esto hace que sea razonable preferir, frente a una teoría de mundos bifurcados, una teoría de mundos divergentes que no se encuentran superpuestos. Entonces debemos ser capaces de hablar de

4 Esta tesis de superveniencia de lo global en lo local se asemeja a las “hipótesis holográficas” consideradas y rechazadas por Saul Kripke en “Identity Through Time”, ensayo presentado en la conferencia de la American Philosophical Association en 1979, División del Este, y en otras partes.

duplicación cualitativa de segmentos de mundo, lo cual puede hacerse en términos de propiedades naturales compartidas.

Mundos divergentes (o bifurcados) son de utilidad a la hora de definir el determinismo. Las definiciones usuales no son muy satisfactorias. Si decimos que cada evento tiene una causa, pasamos por alto la causación probabilística bajo el indeterminismo. Si hablamos de lo que podría ser predicho por una calculadora sobrehumana con conocimiento ilimitado de la historia y las leyes de la naturaleza, pasamos por alto obstáculos que podrían prevenir la predicción incluso bajo determinismo o intentamos dar un sentido no vacuo de los contrafácticos sobre lo que nuestro vaticinador podría hacer si él tuviera una combinación imposible de poderes y limitaciones.

Un mejor enfoque es como el que sigue. Primero, un sistema de leyes de la naturaleza es determinista si y solo si no hay dos mundos divergentes, en donde ambos se conformen perfectamente a las leyes de ese sistema. Segundo, un mundo es determinista si y solo si sus leyes constituyen un sistema determinista. Tercero, el determinismo es la tesis de que nuestro mundo es determinista.⁵

(Versiones alternativas del determinismo pueden ser definidas de manera similar. Por ejemplo, podríamos fortalecer el primer paso al prohibir tanto la convergencia como la divergencia de mundos que se adhieren estrictamente a leyes. O incluso podríamos requerir que no haya dos mundos que adhieran estrictamente a leyes que tengan duplicados de porciones momentáneas sin ser duplicados a lo largo de sus historias. O podríamos definir un tipo más débil de determinismo: llamaríamos a un mundo *fortuitamente* determinista, incluso si sus leyes no constituyen un sistema determinista, si y solo si ningún mundo diverge de él y se conforma a sus leyes. Estas últimas y la historia temprana de dicho mundo son suficientes para determinar su historia posterior, pero solo porque nunca se dan las situaciones en las cuales las leyes no satisfacen el determinismo. Podríamos definir equivalentemente el determinismo fortuito como sigue: por cualquier hecho histórico F y cualquier segmento inicial s del mundo, hay una proposición verdadera H sobre la historia de s y una proposición verdadera L sobre las leyes de la naturaleza, tal que H y L juntas implican estrictamente F .⁶ ¿Acaso

5 Este enfoque se debe, en esencia, a Richard Montague, según lo expuesto en "Deterministic Theories" (1962) y en su *Formal Philosophy* (1974). Pero Montague no habla como yo lo he hecho sobre la duplicación de segmentos iniciales de mundos en virtud del hecho que compartan ciertas propiedades élite. En cambio, utilizó la igualdad de la descripción en un cierto vocabulario, que fue dejado como un parámetro no especificado del análisis, ya que él escribió como un lógico obligado a permanecer neutral sobre preguntas de metafísica.

6 Una definición estrechamente relacionada aparece en Peter Van Inwagen (1975).

esta definición evita nuestra necesidad de hablar sobre la duplicación de los segmentos iniciales? No es así, pues debemos preguntarnos qué quiere decir que *H* es sobre la historia de *s*. Considero que eso significa que *H* vale en ambos o en ninguno de los mundos que comienzan con los segmentos que son duplicados a partir de *s*.)

Los mundos divergentes son también importantes en la conexión con aquel tipo de condicional contrafáctico que figura en los patrones de dependencia causal. Dichos contrafácticos tienden a ser temporalmente asimétricos, y esto es lo que da lugar a la asimetría de la causación en sí misma. Los contrafácticos de este tipo no hacen “marcha atrás”: no debe decirse que si el presente fuese diferente, un pasado distinto nos habría llevado a él, sino más bien que si el presente fuese diferente, el mismo pasado habría tenido un resultado diferente. Dada una diferencia hipotética en un cierto tiempo, los eventos de tiempos futuros normalmente serían, en efecto, muy diferentes, pero los eventos de tiempos pasados (excepto, quizás, por el pasado reciente) no serían diferentes. En consecuencia, la actualidad y sus alternativas contrafácticas son los mundos divergentes, con segmentos iniciales duplicados.⁷

Materialismo mínimo

Existe una dificultad que surge si intentamos formular ciertas visiones reduccionistas, como por ejemplo el materialismo, como una tesis de superveniencia. Una solución a esta dificultad emplea propiedades naturales no solo por medio de la duplicación, sino también de un modo más directo.

A grandes rasgos, el materialismo es la tesis que establece que la física –algo no muy diferente a la física de nuestros días, aunque presumiblemente algo mejorado– es una teoría comprehensiva del mundo, completa y correcta. El mundo es como la física dice que es, y no hay más que decir. La historia del mundo escrita en lenguaje físico es toda la historia del mundo. Esta es una manera tosca de hablar, por cierto; nuestro objetivo será entregar una mejor formulación. Pero antes de intentar decir más precisamente lo que el materialismo es, déjeme decir lo que no es. (1) El materialismo no es la tesis de la traducibilidad finita de todo nuestro lenguaje al lenguaje de la física. (2) El materialismo no debe ser identificado con ninguna teoría materialista de la mente. Es una tesis que motiva una variedad de teorías de la mente: versiones del conductismo, el funcionalismo, la teoría de la identidad mente-cuerpo, incluso la teoría de que la mente que es un completo error. (3) El materialismo no es solo la teoría de que no hay cosas, excepto aquellas

7 Véase mi “Counterfactual Dependence and Time’s Arrow” (1979); la reseña de Jonathan Bennett de mi *Counterfactuals* (1974); Downing (1959); Allan Gibbard y William Harper (1978), y Frank Jackson (1977).

reconocidas por la física. Para estar seguros, los materialistas no creen en espíritus u otras cosas no-físicas. Pero los antimaterialistas podrían no creer tampoco en espíritus –no es necesario que su queja sea que la física omita algunas de las cosas que hay–. Pueden alegar, en cambio, que la física pasa por alto algunas de las maneras en que las cosas físicas difieren; por ejemplo, podrían pensar que las personas físicas pueden diferir con respecto a cómo es su experiencia. (4) Eso sugiere que el materialismo es, por lo menos en parte, la tesis de que no hay propiedades naturales instanciadas en nuestro mundo excepto aquellas reconocidas por la física. Eso está mejor pero no creo que sea correcto. ¿Acaso no podría haber una propiedad natural x (dada la naturaleza del caso, ¡es muy difícil dar un ejemplo!) que sea compartida por los cerebros físicos en los mundos como el nuestro y los espíritus inmateriales que habitan otros mundos?, ¿o por quarks de este mundo y ciertas partículas de otro que no existan bajo nuestra física? La física podría perfectamente no hacer mención de una propiedad natural de este tipo. Es suficiente con reconocer el caso especial aplicable a nuestro mundo, tal como podría ser *x-cum-fisicalidad*, cerebriedad (*brainhood*), quarkedad (*quarkhood*). Entonces, si por propiedades físicas nos referimos a aquellas que son mencionadas en el lenguaje de la física, un materialista no debe sostener que todas las propiedades naturales instanciadas en nuestro mundo son propiedades físicas.

En este punto, debería ser recomendable formular el materialismo como una tesis de superveniencia: no hay diferencia sin diferencia física. O, por contraposición, los duplicados físicos son duplicados *simpliciter*. *A fortiori*, no hay diferencia mental sin distinción física; los duplicados físicos son duplicados mentales. La tesis puede ser mejor entendida si se aplica a mundos posibles completos, para así evitar preguntas como la de si acaso la vida mental es, bajo algún respecto, extrínseca al sujeto. Es así como tenemos el primero de varios intentos de formulación del materialismo:

M1 Dos mundos posibles, cualesquiera que sean, exactamente iguales bajo todos los aspectos reconocidos por la física son duplicados cualitativos.

Pero esto no es suficiente. Al convertir el materialismo en una tesis sobre cómo dos mundos cualesquiera pueden o no diferir, *M1* posiciona esta teoría como una verdad necesaria. Esa no es la intención de los materialistas. El materialismo está destinado a ser una tesis contingente, un mérito de nuestro mundo que no es compartido por todos los demás. Dos mundos podrían, en efecto, diferir sin distinguirse físicamente, si a lo menos uno de ellos es un ámbito en donde el materialismo es falso. Por ejemplo, nuestro mundo materialista difiere de uno no-materialista

que es físicamente como el nuestro pero que también contiene espíritus físicamente epifenoménicos.

Hay una tesis cercana de superveniencia no-contingente que podría interesar a los materialistas:

M2 No hay diferencia, *a fortiori* no hay diferencia mental, sin alguna distinción no-mental. Dos mundos cualesquiera iguales bajo todos los aspectos no-mentales son duplicados, y en particular no difieren respecto de las vidas mentales de sus habitantes.

Esto parece aprehender nuestro pensamiento de que lo mental es un patrón en un medio, que se obtiene en virtud de características locales del medio (disparos neuronales) y quizá también de características muy globales (leyes de la naturaleza) que son demasiado pequeñas o demasiado grandes para ser mentales en sí mismas. Pero *M2* no es materialismo. Es, al mismo tiempo, algo más y algo menos. Menos, obviamente, porque nunca dice que el medio es físico; más porque niega la posibilidad de lo que llamaré materialismo *pampsíquico* (*panpsychistic materialism*).

Se destaca habitualmente que la identidad psicofísica es una calle con dos sentidos: si todas las propiedades mentales son físicas, entonces algunas de estas son mentales. Pero quizá no algunas sino *todas* las propiedades físicas podrían ser mentales; y, en efecto, cada propiedad de cualquier cosa podría ser al mismo tiempo mental y física. Supóngase que, en efecto, hay mundos en donde esto es así. Entonces, presumiblemente hay muchos mundos de esta manera, no todos duplicados, que difieren *inter alia* en las vidas mentales de sus habitantes. Pero todas las diferencias entre dichos mundos son mentales (así como también físicas), por lo que ninguna es no-mental. Estos mundos serán vacuamente iguales bajo todos los aspectos no-mentales, a falta de cualquier aspecto no-mental según el cual puedan diferir. Entonces, *M2* fracasa, y no solo en los mundos que causan problemas; *M2* es no-contingente, por lo que si falla en cualquiera de los mundos, entonces lo hace en todos –incluso en los mundos materialistas decentes como el nuestro-. Quizás el materialismo pampsíquico es imposible –¿cómo es posible hacerlo calzar con un análisis general funcional de la mente?– pero una tesis que lo afirme es mucho más que el materialismo.

Un tercer intento. Este por lo menos es verdadero:

M3 No hay dos mundos materialistas que difieran sin distinguirse físicamente; dos mundos materialistas cualesquiera que sean exactamente iguales físicamente son duplicados.

Pero *M3* no es una formulación del materialismo, ya que la distinción entre mundos materialistas y otros aparece dentro de *M3*. Todo lo que aprendemos es que los mundos materialistas incluyen una clase dentro

de la cual no hay diferencia sin diferencia física. Pero hay muchas clases así. De hecho, cualquier mundo, por muy atestado de espíritus que se encuentre, pertenece a dicha clase.

Un cuarto intento. Quizá deberíamos centrar nuestra atención en mundos posibles nomológicos, así:

M4 Entre los mundos que se ajustan a las leyes de la naturaleza actuales, no hay dos que difieran sin hacerlo físicamente; dos mundos cualesquiera de ese tipo que sean exactamente iguales físicamente son duplicados

Pero nuevamente tenemos algo que es, al mismo tiempo, algo más que el materialismo y algo menos. Menos, porque *M4* podría darse en un mundo en donde el materialismo es falso, pero en el que los fenómenos espirituales están correlacionados con fenómenos físicos según leyes estrictas. Más, porque *M4* no se da en un mundo materialista libre de espíritus si las leyes de este no excluyen la existencia de espíritus epifenómicos. Nuestro mundo podría ser ese, uno en donde los espíritus están ausentes pero no proscritos.⁸

Hasta el momento, una formulación del materialismo como superveniencia parece esquiva. Pero creo que podemos tener éxito si unimos la idea de superveniencia con la de que un mundo no-materialista pueda tener algo extra, algo que un mundo materialista carece. Podría tener espíritus o podría tener cosas físicas que difieren en aspectos no-físicos, por ejemplo, en cómo es su experiencia. En cualquier caso, hay propiedades naturales adicionales, propiedades instanciadas en el mundo no-materialista, pero sin estar en ninguna parte del mundo materialista. Digamos que una propiedad es *extranjera* (*alien*) a un mundo si y solo si (1) no está instanciada por ningún habitante de este, y (2) no es analizable como una conjunción de, o como una propiedad estructural construida a partir de, propiedades naturales que se encuentran todas ellas instanciadas por habitantes de ese mundo. (Necesito la segunda cláusula porque estoy siguiendo a Armstrong, *mutatis mutandis*, al negarme a dejar por fuera a las propiedades perfectamente naturales que son conjuntiva o estructuralmente complejas. Véase Armstrong 1978 II: 30-42 y 67-71. Sería un error contar como extranjera una propiedad compleja analizable en términos de constituyentes no-extranjeras.) Si nuestro mundo es materialista, entonces es seguro decir que algunas de las propiedades naturales instanciadas en cualquier mundo no-materialista son propiedades extranjeras con relación a nuestro mundo. Ahora podemos finalmente proceder a formular el materialismo como una tesis restringida y contingente de superveniencia:

8 Esta objeción contra *M4* como una formulación de “la primacía ontológica de lo microfísico” aparece en Terence Horgan (1982).

M5 Entre mundos donde hay instanciadas propiedades que no son extranjeras a nuestro mundo, no hay dos que difieran sin hacerlo físicamente; dos mundos cualesquiera de ese tipo que sean exactamente iguales son duplicados.⁹

★

Hemos considerado que el materialismo defiende la exhaustividad de “algo no tan diferente de la física actual, aunque presumiblemente algo mejorada”. Eso fue deliberadamente vago. Los metafísicos materialistas quieren alinearse con la física, pero no tomar una posición dentro de esta última. Dentro de la física, pretensiones más precisas sobre exhaustividad y exactitud pueden estar en cuestión. La física (ignorando el desaliento de los últimos tiempos) es la ciencia que aspira a la exhaustividad, y las teorías físicas particulares pueden ser consideradas como maneras de cumplir esa aspiración. Si es así, entonces debemos preguntarnos nuevamente qué significa alegar exhaustividad. Nuevamente, la respuesta puede ser dada por una formulación de superveniencia: no hay diferencia sin distinción física, según como sea concebida por tal-y-tal gran teoría. Pero nuevamente debe ser entendida como una tesis de superveniencia restringida y contingente, aplicable solo a los mundos desprovistos de propiedades naturales extranjeras.

En consecuencia, la tarea de la física no consiste solamente en descubrir leyes y explicaciones causales. Al presentar teorías exhaustivas que reconocen solo un limitado rango de propiedades naturales, la física propone inventarios de propiedades naturales instanciadas en nuestro mundo; tal vez no inventarios completos pero sí lo suficientemente completos como para dar cuenta de todas las duplicaciones y diferencias que podrían surgir ante la ausencia de propiedades naturales extranjeras. Por supuesto, el descubrimiento de propiedades naturales es inseparable del hallazgo de leyes, ya que una excelente razón para pensar que algunas propiedades naturales insospechadas hasta ahora

9 Esta formulación se asemeja a una propuesta por Horgan (1982). La diferencia principal es la que sigue. Este autor contaría como extranjera (mi término, no el suyo) cualquier propiedad citada en las leyes fundamentales de la microfísica de otros mundos que no se encuentre, a su vez, citada explícitamente en las leyes fundamentales de la microfísica de este mundo. El que la propiedad esté instanciada en cualquiera de los mundos no es considerado. ¿Pero debe una propiedad extranjera figurar en las leyes de la física de otro mundo? ¿Debe siquiera figurar en las leyes de otro mundo? Pareciera ser que un mundo materialista podría diferir sin hacerlo físicamente de uno en donde hay propiedades extranjeras en mi sentido pero no en el de Horgan –quizás un mundo en donde las leyes son pocas–.

están instanciadas –propiedades que merezcan el reconocimiento de la física, como podrían ser los colores de quark– es que sin ellas ningún sistema satisfactorio de leyes puede ser encontrado.

Esto es una reminiscencia del distintivo carácter científico *a posteriori* del realismo de Armstrong (cf. 1978 I: 8-9 y *passim*). Pero en el escenario de una ontología de *possibilia* se desvanece la distinción entre el descubrimiento de los universales o propiedades naturales que en realidad hay y el descubrimiento de aquellas que están realmente instanciadas. Y esta última pregunta se hace *a posteriori* en cualquier teoría. Lo que queda, y permanece como importante, es que la física descubre propiedades, y no cualquier tipo de ellas –propiedades naturales–. El descubrimiento es, por ejemplo, que no todos los neutrinos son iguales. Eso no es el descubrimiento de que neutrinos diferentes tienen diferentes propiedades según mi sentido, perteneciendo a diferentes clases. Puesto que ya sabíamos eso *a priori*; más bien, es el sorprendente descubrimiento de que alguna propiedad *natural* diferencia algunos neutrinos de otros. Ese descubrimiento de hecho se ha efectuado; me gustaría leer a un filósofo que dé cuenta de esto que no esté preparado para adoptar una actitud discriminadora hacia las propiedades y que piense que todas las cosas son igualmente similares y no similares unas con otras.

Leyes y causación

La observación acerca de que la física descubre propiedades en el curso del hallazgo de leyes puede servir para introducir nuestro siguiente tema: el análisis de qué es lo que significa ser una ley de la naturaleza. Conuerdo con Armstrong en que necesitamos universales, o a lo menos propiedades naturales, para explicar lo que una ley es, aunque estoy en desacuerdo con su explicación sobre cómo debe ser.

La teoría de Armstrong, en su forma más simple,¹⁰ sostiene que lo que hace que ciertas regularidades tengan el carácter de ley es la existencia de estados de cosas de segundo orden $N(F, G)$, según la cual los dos universales ordinarios de primer orden F y G se encuentran relacionados por un cierto universal diádico de segundo orden, N . Es un asunto contingente cuáles universales están relacionados de esta manera con el hacedor de leyes N . Pero es necesario –y necesario *simpliciter*, no solo nomológicamente necesario– que si $N(F, G)$ se da, entonces F y G están constantemente unidos. Hay una conexión necesaria entre el estado de cosas de segundo orden $N(F, G)$ y la regularidad con carácter de ley de

10 Armstrong (1978 II: 148-157). Una forma más desarrollada de la teoría aparece en Armstrong (1983). Teorías similares han sido propuestas por Fred I. Dretske (1977) y Michael Tooley (1977).

primer orden $\forall x (Fx \supset Gx)$ y asimismo entre el estado de cosas conjuntivo $N(F, G)$ y Fa y su consecuencia necesaria Ga .

Una teoría paralela podría ser montada con propiedades naturales en lugar de los universales de primer y segundo orden de Armstrong. Tendría muchas de las atractivas características que este autor predica de su teoría, pero por lo menos un mérito se perdería. Para Armstrong la necesitación (*necessitation*) legal de Ga por Fa es un asunto puramente local: involucra solamente a a , a los universales F y G que están presentes en a y al universal hacedor de leyes de segundo orden que está presente en (o entre) estos dos universales. Si reemplazamos a los universales por propiedades, por muy naturales que sean, la localidad se pierde, ya que las propiedades son clases cuyos miembros están esparcidos por los mundos y no están completamente presentes en a . Pero no creo que esta sea una objeción concluyente, pues nuestras intuiciones sobre la localidad suelen extraviarnos. La teoría de la regularidad selectiva que pronto defenderé también sacrifica la localidad, como lo hace cualquier teoría de regularidad de las leyes.

Lo que me lleva (con cierto pesar) a rechazar la teoría de Armstrong, ya sea con universales o con propiedades naturales, es que considero que sus conexiones necesarias son ininteligibles. Sea lo que fuere N , no puedo ver cómo podría ser absolutamente imposible tener $N(F, G)$ y Fa sin Ga (a menos que N justamente sea una conjunción constante o una conjunción contante con algo más, en cuyo caso la teoría de Armstrong se convierte en una forma de teoría de regularidad que él rechaza). El misterio está algo escondido en la terminología de este autor. Él utiliza “necesita” (*necessitates*) como un nombre para el universal hacedor de leyes N ; ¿y a quién le sorprendería escuchar que si F “necesita” a G y a tiene F , entonces a debe tener G ? Pero sostengo que N merece el nombre de “necesitación” solo si, de alguna manera, puede ingresar en las conexiones necesarias indispensables. No puede ingresar en ellas solamente por ostentar el nombre, más de lo que uno puede tener bíceps poderosos solo por llamarse “Armstrong”. **

Estoy tentado quejarme, en un estilo humeano, de las presuntas conexiones necesarias entre existencias distintas, especialmente cuando los estados de cosas de primer orden del pasado supuestamente se unen con la estados de cosas de segundo orden para necesitar estados de cosas de primer orden en el futuro. Esta queja no es claramente correcta: el hecho de compartir universales le quita valor a la distinción entre estados de cosas que necesitan y estados de cosas necesitados. Pero no estoy apaciguado. Concluyo que las conexiones necesarias pueden ser

.....
 **Nota del traductor: David Lewis está empleando un juego de palabras, pues el apellido Armstrong resulta de la unión de “arm” y “strong”, i.e. “brazo” (o bíceps) y “poderoso”.

ininteligibles aun cuando supuestamente son obtenidas entre existencias que no son clara y completamente distintas.¹¹

En consecuencia, no apruebo la manera en que Armstrong introduce los universales, o alternativamente las propiedades naturales, en el análisis de la legalidad. En su lugar, favorezco un análisis de regularidad. Pero incluso para esto requiero de propiedades naturales.

Ciertamente no cualquier regularidad es una ley de la naturaleza. Algunas son accidentales. Por lo tanto, un adecuado análisis de la regularidad debe ser selectivo; a su vez, este mismo debe ser colectivo. Este no debe tratar las regularidades una a la vez, sino más bien como candidatas para ingresar a sistemas integrados, ya que una regularidad dada pueda darse tanto como una ley o accidentalmente, dependiendo de si acaso otras regularidades pueden calzar con ella en un sistema adecuado (por lo tanto, rechazo la idea de que la legalidad consiste en “un parecido de ley” más la verdad). Siguiendo a Mill y a Ramsey,¹² considero que un sistema adecuado es aquel que tiene las virtudes a las que aspiramos en nuestra propia construcción de teorías, y que las tiene en la mayor extensión posible, dada la manera en que el mundo es. Debe ser completamente verdadera, debe ser estar cerrada bajo la implicación estricta, debe ser lo más simple posible en axiomatización sin sacrificar mucho contenido informativo y debe tener la mayor cantidad de contenido informativo sin sacrificar mucha simplicidad. Una ley es cualquier regularidad que gane su inclusión en un sistema ideal (o, en caso de empate, en cada sistema ideal). No es necesario que el sistema

11 La teoría más desarrolla de Armstrong en *What is a Law of Nature?* complica el panorama de dos maneras. Primero, el estado de cosas de segundo orden $N(F, G)$ es considerado en sí mismo como un universal, y su presencia en sus instancias le quita más valor aún a la distinción entre estados de cosas que necesitan y estados necesitados. Segundo, todas las leyes son anulables. Es posible, después de todo, tener $N(F, G)$ y Fa sin tener Ga , a saber, si también tenemos $N(E \vee F, H)$ y Ea , donde H y G son incompatibles. La ley que establece que los F son G podría ser *contingentemente* no-anulable, si no se da el estado anulable $N(E \vee F, H)$; pero ninguna ley tiene su no-anulabilidad incluida en sí esencialmente. Permanece verdadero que hay presuntas conexiones necesarias que considero ininteligibles, pero son más complicadas que antes. Para necesitar un estado de cosas, no solo requerimos los estados de cosas de primer y segundo orden originalmente considerados, sino también un existencial negativo en el sentido de que no haya más estados de cosas del tipo que puedan actuar como anuladores.

12 John Stuart Mill *A System of Logic*, Libro III, Capítulo IV, Sección 1 (1843); Ramsey “Universals of Law and of Fact”, en su *Foundations* (1978). Ramsey consideraba que esta teoría de la ley estaba sustituida por una teoría diferente presentada en su “General Propositions and Causality”, también en su *Foundations*, pero prefiero sus primeros pensamientos a los segundos. Yo presento una teoría de la legalidad siguiendo la línea de la teoría temprana de Ramsey en mi *Counterfactuals* (cf. 1973 73-75). Una revisión de este caso es necesaria en el caso probabilístico, el cual ignoro aquí.

ideal consista enteramente en regularidades; los hechos particulares pueden ganar acceso si es que contribuyen lo suficiente a la simplicidad y fuerza colectivas (por ejemplo, ciertos hechos particulares sobre el Big Bang podrían ser candidatos muy fuertes). Pero solo las regularidades del sistema pueden contar como leyes.

Nos enfrentamos a un problema obvio. Diferentes maneras de expresar el mismo contenido, utilizando distintos vocabularios, diferirán en simplicidad. El problema puede ser presentado de dos maneras, dependiendo de si acaso consideramos que nuestros sistemas están compuestos de proposiciones (clases de mundos) o de oraciones interpretadas. En el primer caso, el problema es que un mismo sistema tiene diferentes grados de simplicidad relativos a diferentes formulaciones lingüísticas. En el segundo, el problema es que sistemas equivalentes, que implican estrictamente las mismas regularidades, pueden diferir en su simplicidad. De hecho, el contenido de cualquier sistema puede ser formulado de manera muy simple. Dado el sistema s , sea F un predicado que se aplica a todas y solo las cosas en los mundos en donde s se da. Tómese F como primitivo y axiomatícese S (o algún equivalente de s) por un único axioma $\forall xFx$. Si la simplicidad extrema puede ser tan fácilmente obtenida, la teoría ideal, del mismo modo, puede ser tan fuerte como sea posible. No es necesario que la simplicidad y la fuerza sean compensadas entre sí. Entonces la teoría ideal incluirá todas las verdades (su axioma simple las implicará estrictamente) y *a fortiori* todas las regularidades. Entonces, después de todo, cada regularidad será una ley. Eso debe ser incorrecto.

El remedio, por supuesto, está en no tolerar una elección tan perversa de vocabulario primitivo. Deberíamos preguntar cómo los candidatos a sistemas se comparan en simplicidad cuando cada uno está formulado de la manera más simple; o, si es que contamos diferentes formulaciones como diferentes sistemas, deberíamos descartar aquellos que no sean aptos para ser candidatos. No hay que buscar lejos un estándar apropiado de elegibilidad: sea que el vocabulario primitivo que aparece en los axiomas se refiera solamente a propiedades perfectamente naturales.

Por supuesto, sigue siendo un problema difícil y no resuelto determinar lo que la simplicidad de una formulación es. Pero ya no es la cuestión derechamente insoluble que sería si es que no pudiéramos escoger entre alternativas de vocabularios primitivos.

(Uno podría pensar en reemplazar la implicación estricta por la deducibilidad en algún cálculo específico. Pero este segundo remedio no parece necesario dado el primero y parece incapaz de resolver nuestro problema por sí solo.)

Si adoptamos el remedio propuesto, tendremos que las leyes tenderán a ser regularidades que involucren propiedades naturales. Las leyes

fundamentales, aquellas que el sistema ideal considera como axiomas, deben ser concernientes a propiedades perfectamente naturales. Las leyes derivadas que se siguen de modo bastante directo también tenderán a ser concernientes a propiedades bastante naturales. Las regularidades que sean concernientes a propiedades no-naturales pueden ser, en efecto, implicadas estrictamente y deberían ser contadas como leyes derivadas si es así. Pero ellas son aptas para eludir nuestra atención, incluso si es que algún día logramos poseer una buena aproximación a un sistema ideal, ya que sería difícil expresarlas en un lenguaje que tenga palabras para propiedades no tan no-naturales, como debería ser cualquier lenguaje (véase la siguiente sección), y serían difíciles de derivar, incluso podrían no ser finitamente derivables del todo, en nuestro cálculo deductivo. En consecuencia, mi propuesta explica, como la de Armstrong lo hace de una manera muy diferente, por qué la investigación científica de las leyes y las propiedades naturales es un paquete completo; por qué los físicos postulan propiedades naturales tales como los colores de quarks para poder postular leyes en las cuales figuren esas propiedades, de tal manera que las leyes y las propiedades se descubren al mismo tiempo.

*

Si el análisis de la legalidad requiere de propiedades naturales, entonces también las precisa el análisis de la causación. Es poco controversial sostener que la causación involucra leyes. Esto es así de acuerdo con las dos principales teorías de la causación: el análisis nomológico-deductivo, según el cual las leyes se aplican al curso actual de los eventos con la causa y el efecto presentes, y el análisis contrafáctico que yo favorezco, según el cual las leyes se aplican a situaciones contrafácticas con la causa como una hipótesis. Estas alternativas contrafácticas pueden llegar a romper las leyes actuales en el punto en donde divergen de la actualidad, pero el análisis requiere que evolucionen después de eso acorde a las leyes actuales.¹³

Según mi análisis contrafáctico, la causación involucra también de una segunda manera propiedades naturales. Necesitamos el tipo de contrafácticos que eviten la marcha atrás; de otra manera, el análisis se enfrenta a contraejemplos fatales que involucran efectos secundarios epifenómicos o casos de prevención causal (*causal preemption*). Como ya lo he mostrado, estos contrafácticos deben ser caracterizados en términos de mundos divergentes, o sea en términos de duplicados de segmentos iniciales de mundo, esto es, en términos de propiedades naturales compartidas.

13 Véase mi "Causation" (1973; reimpreso en 1975).

La causación involucra propiedades naturales, incluso de otra manera más (¡no me sorprende que haya llegado a apreciar las propiedades naturales después de trabajar en el análisis de la causación!). Esta última se da entre eventos. A menos que distingamos entre eventos genuinos y espurios, obtendremos muchas causas putativas. Usted pone un pedazo de mantequilla en un sartén, y aquella se derrite. ¿Qué evento causa esto? Hay un evento que podemos llamar el movimiento de las moléculas. Ocurre en la región en donde está el sartén, antes de que la mantequilla se derrita. Este es un evento tal que, necesariamente, ocurre en una región espacio-temporal solo si esta contiene moléculas que se mueven rápidamente. Seguramente este evento es la causa del derretimiento de la mantequilla.

El calor es un fenómeno, sea lo que sea, que se manifiesta a sí mismo de cierta manera familiar y característica. Digamos: el calor es aquello que ocupa el papel-del-calor (*heat-role*) (no importa si acaso tomamos la descripción definitiva simple, o como yo prefiero, rigidizada); de hecho, pero contingentemente, es el movimiento molecular el que ocupa el papel-del-calor. Podría haber sido el no-movimiento molecular, o el fluido calórico, o lo que se quiera. Ahora considérese un segundo presunto evento, uno que podemos llamar poseer-el-ocupante-del-papel-del-calor. Este segundo evento ocurre donde y cuando lo hace el primero, en la región en donde está el sartén caliente. Ocurre en virtud de dos hechos (1) que las moléculas del sartén están moviéndose rápidamente y (2) que la región en cuestión es parte de un mundo en donde el movimiento molecular es lo que ocupa el papel del calor. Pero este segundo evento difiere del primero. Las condiciones necesarias para su ocurrencia son distintas. Necesariamente, se da en una región solo si esta contiene el fenómeno, cualquiera que este sea, que ocupa el papel del calor en el mundo del cual la región es parte. Entonces, en aquellos mundos en donde el fluido calórico ocupa el papel del calor y el movimiento molecular no, el primer evento ocurre solo en las regiones con movimiento molecular, mientras que el segundo solo ocurre en aquellas regiones con fluido calórico.

Ciertamente, el primer evento causa el derretimiento de la mantequilla, pero ¿debemos decir que el segundo evento también lo hace? No; esto parece multiplicar las causas más allá de la creencia utilizando un truco verbal. Pero si realmente hay dos eventos aquí, no veo por qué el segundo tiene menos derecho que el primero para ser la causa del derretimiento de la mantequilla. Está fuera de discusión que el primer y el segundo evento son uno y el mismo –pues este único evento tendría condiciones de ocurrencia diferentes de sí mismo–. La mejor solución es negar que el presunto segundo evento sea un evento genuino del todo. Si no lo es, por supuesto que no puede causar nada.

¿Por qué es el primer evento genuino y el segundo espurio? Compárese las propiedades involucradas: contener moléculas que se mueven rápidamente frente a contener un fenómeno, cualquiera que sea, que ocupe el papel de calor (me refiero a estas como propiedades de la región espacio-temporal; otros tratamientos de los eventos considerarían las propiedades correspondientes al sartén, pero mi punto se aplicaría de todos modos). La primera es una propiedad intrínseca y bastante natural; la segunda es altamente disyuntiva y extrínseca, ya que una gran cantidad de fenómenos diferentes podrían ocupar el papel de calor; y si acaso el fenómeno que está ocurriendo en la región ocupa este papel depende no solo de lo que ocurra en la región, sino también de lo que esté sucediendo en el resto del mundo. Es así como la distinción entre propiedades más o menos naturales me da la distinción entre eventos genuinos y espurios que necesito para repudiar la sobreabundancia de causas. Si una propiedad es muy poco natural, es ineficaz en el sentido de que no puede figurar en las condiciones de ocurrencia de los eventos que causan cosas.¹⁴

El contenido del lenguaje y el pensamiento

Hilary Putnam ha dado un argumento que considera como una refutación de una visión “radicalmente no-epistémica” de la verdad, pero que yo considero más bien como una *reductio* contra las premisas de Putnam.¹⁵ En particular, refuta su suposición de que “o bien *nosotros* interpretamos nuestro lenguaje o nada lo hace” (Putnam 1980 482) por lo que cualquier restricción en la referencia debe ser establecida por nuestra estipulación en el lenguaje o en el pensamiento. Gary Merrill ha sugerido que a Putnam se le puede contestar apelando a una restricción que dependa de una estructura objetiva de propiedades y relaciones del mundo.¹⁶ Estoy de acuerdo, y aquí encuentro otro punto en donde necesitamos propiedades naturales.

14 Véase la discusión de la impotencia de las disposiciones en Elizabeth W. Prior, Robert Pargetter y Frank (1982). Si una disposición no está identificada con su base actual, hay una amenaza de multiplicación de causas putativas similar a aquella en mi ejemplo. No quisiéramos decir que la ruptura de un vidrio golpeado es causada tanto por su fragilidad como por las tensiones residuales que son la base de la fragilidad; y si nos fuerzan a elegir, deberíamos optar por lo último. Sugiero que la fragilidad es ineficaz porque es una propiedad muy poco natural, demasiado disyuntiva y extrínseca, para figurar en las condiciones de ocurrencia de cualquier evento.

15 Hilary Putnam (1978, 1980). El lector debe estar advertido de que el argumento que presento puede que no sea tal como Putnam lo ha pretendido, ya que he hecho una lectura libre entre líneas y he vuelto a exponer el argumento a mi modo.

16 Merrill (1980).

El argumento de Putnam, como yo lo entiendo, es como sigue. Primero, supóngase que la única restricción sobre la interpretación de nuestro lenguaje (o quizá nuestro lenguaje del pensamiento) está dada por una teoría descriptiva de la referencia de un tipo global y futurista. Una “interpretación pretendida” es una interpretación que satisface un cierto cuerpo de teoría: *viz.* el descendiente idealizado de nuestra teoría total presente que emergería al final de la investigación, una teoría ideal refinada a la perfección bajo la guía de toda la observación necesaria y nuestro mejor razonamiento teórico. Si esto es así, las interpretaciones pretendidas son sorprendentemente abundantes, ya que *cualquier* mundo puede satisfacer *cualquier* teoría (ideal o no) y puede hacerlo de innumerables maneras, concediéndose solamente que el mundo no sea demasiado pequeño y la teoría sea consistente. Más allá de eso, no importa cómo es el mundo o qué dice la teoría. Por consiguiente, tenemos una radical indeterminación de la referencia y la coincidencia que Putnam acepta gustoso entre todas las interpretaciones pretendidas y la “verdad epistémica”. Ya que la teoría ideal es la “verdad epistémica” completa, las interpretaciones pretendidas son solo aquellas interpretaciones de nuestro lenguaje que satisfacen la teoría ideal, y (a menos que el mundo sea demasiado pequeño o la teoría inconsistente), tales interpretaciones existen.

Considero que esto refuta la suposición de que no hay restricciones adicionales sobre la referencia. Pero Putnam pregunta: ¿cómo *podría* haber una restricción adicional? ¿Cómo podríamos alguna vez establecerla? Por estipulación, al decir o pensar algo. Pero cualquier cosa que digamos o pensemos estará formulada en un lenguaje (o el lenguaje del pensamiento) que sufre una radical indeterminación de interpretación. Ya que la restricción salvadora no estará ahí hasta que la establezcamos exitosamente. Por lo tanto, la estipulación intentada debe fracasar. Lo mejor que podemos hacer es contribuir con un nuevo capítulo a la teoría actual e ideal, uno consistente en lo que sea que hayamos dicho o pensado en nuestra estipulación. Esta nueva teoría sigue el camino de toda teoría, por lo que no podemos establecer una restricción adicional; o bien “*nosotros* interpretamos nuestro lenguaje o nada lo hace”; luego, no puede haber una restricción adicional. Nosotros no podemos elevarnos tirando de nuestras propias botas, por lo que debemos seguir en el suelo.

En efecto, no podemos elevarnos de esta manera, pero hemos despegado y estamos sobre el suelo; por consiguiente, debe haber otra manera de volar. Nuestro lenguaje tiene una cierta interpretación determinada (¡un hecho mooreano!), por lo que debe haber alguna restricción que no haya sido creada *ex nihilo* por nuestra estipulación.

¿Qué puede ser? Muchos filósofos sugerirían que se trata de algún tipo de restricción causal. Si es así, entonces mi tesis está armada, dados

mis argumentos en la sección anterior: necesitamos propiedades naturales para explicar la determinación de la interpretación. Pero dudo si verdaderamente estamos frente a una restricción causal, pues estoy inclinado a pensar que el aspecto causal de la referencia *está* establecido por lo que decimos y pensamos. En consecuencia, yo pienso en una cosa como aquella con la que estoy causalmente familiarizado de tal-y-tal manera, quizá perceptivamente o quizás a través de un canal de familiaridad (*acquaintance*) que involucra el nombramiento de la cosa y mi elección del nombre. Yo me refiero a esa cosa en mi pensamiento y de manera derivada en el lenguaje, porque es ella la que calza con esta descripción causal y egocéntrica que ha sido extraída de mi teoría sobre el mundo y sobre mi lugar en el mundo.¹⁷

Propondría, en cambio, que la restricción salvadora concierne al referente –no al que refiere ni a los canales causales entre ambos–. Se requieren dos para hacer la referencia, y no encontraremos la restricción si siempre la buscamos en el lado equivocado de la relación. La referencia consiste, en parte, en lo que hacemos en el lenguaje o en el pensamiento cuando referimos, pero en parte se trata de la elegibilidad del referente, y esta, para ser referida, es un asunto de propiedades naturales.

Esa es la sugerencia que Merrill ofrece (no la ofrece como su propia visión, sino como lo que los opositores de Putnam deberían decir; y yo agradecidamente acepto el ofrecimiento). En el caso más simple, supóngase que la interpretación del vocabulario lógico de alguna manera se encarga de sí misma, para revelar un lenguaje estándar de primer orden, cuyo vocabulario no-lógico consiste enteramente de predicados. Las partes del mundo incluyen un dominio, y los conjuntos, conjuntos de pares, etc., de este dominio son las extensiones potenciales para los predicados. Ahora supóngase que tenemos una división de todo-o-nada entre propiedades naturales y no naturales. Digamos que un conjunto del dominio es *elegible* para ser la extensión de un predicado con una variable (*one-place predicate*) si y solo si sus miembros son aquellas cosas en el dominio que comparten alguna propiedad natural y, del mismo modo, para predicados de más de una variable (*many-place predicate*) y relaciones naturales. Una *interpretación elegible* es aquella que asigna solo extensiones elegibles a los predicados. La llamada interpretación “pretendida” es una interpretación elegible que satisface la teoría ideal (pero el nombre es engañoso: no debe decirse que nuestras intenciones establecen la restricción que requiere la elegibilidad. Por ese camino está el fútil intento, que debemos evitar, de levantarnos tirando de nuestras propias botas). Entonces, si las propiedades naturales son escasas, no hay

17 Véase Stephen Schiffer (1978).

razón para esperar una sobreabundancia de interpretaciones pretendidas, incluso puede que no haya ninguna, y la teoría ideal corre el riesgo de no ser satisfecha, salvo de maneras “no pretendidas”. Dado que la satisfacción no está asegurada, conseguiremos algo si logramos alcanzarla al hacer calzar la teoría y el mundo. Todo esto es como debe ser.

La propuesta requiere refinamiento. Primero, necesitamos proporcionar formas más ricas de lenguaje. En esto podemos ser guiados por traducciones familiares, por ejemplo, entre el lenguaje modal con cuantificación de orden superior y el lenguaje de primer orden que explícitamente menciona *possibilia* y clases construidas sobre ellos. Segundo, de nada sirve considerar la naturalidad de las propiedades como todo-o-nada. Aquí, sobre todas las cosas, necesitamos hacer de la naturalidad –y, en consecuencia, de la elegibilidad– un asunto comparativo o de grados. Hay líneas precisas y salientes, pero no en los lugares correctos. Está la línea entre las propiedades perfectamente naturales y las demás, pero seguramente tenemos predicados para propiedades menos que perfectamente naturales; está la línea entre las propiedades que son y no son finitamente analizables en términos de propiedades perfectamente naturales, pero esto deja entrar propiedades altamente no-naturales en cantidades suficientes como para poner en riesgo la solución a nuestro problema. Necesitamos gradaciones y un poco de tira y afloja entre la elegibilidad de los referentes y otros factores que tengan incidencia en el carácter de “pretendido”, en especial la satisfacción de partes apropiadas de teoría (teoría ideal, si retenemos lo más posible del relato de Putnam). Lo verdul no es un referente absolutamente inelegible (como justo ahora lo muestra mi referencia a aquello), pero una interpretación que lo asigna es en aquella medida inferior que una que asigna azulidad. *Ceteris paribus*, la última es la “pretendida” simplemente porque le va mejor en términos de elegibilidad.

La naturalidad de las propiedades hace que haya diferencias de elegibilidad no solo entre las propiedades mismas, sino también entre las cosas. Compárese a Bruce con un pedazo de materia miscelánea y altamente cambiante, con forma de gato, que lo sigue a todas partes, siempre unos pasos más atrás. El primero es un referente altamente elegible, el segundo no lo es (no he tenido éxito al referirme a eso, ya que no dije qué pedazo “eso” debía ser). Eso es porque Bruce, a diferencia del pedazo de materia con forma de gato, tiene una frontera bien demarcada por propiedades altamente naturales. Donde Bruce termina, la densidad de la materia, la abundancia de elementos químicos, etc., abruptamente cambia. No así para el pedazo de materia. Bruce también es mucho más el lugar para cadenas causales que el pedazo de materia; esto también se puede rastrear en las propiedades naturales, según las consideraciones de la sección anterior. Así, la naturalidad de las

propiedades establece distinciones entre las cosas. Lo inverso también sucede. Una vez que estamos lejos de las propiedades perfectamente naturales, algo que provoca la naturalidad de una propiedad es que esta pertenece exclusivamente a cosas bien demarcadas.

*

Uno podría protestar diciendo que el problema de Putnam está mal entendido, razón por la cual no se ha demostrado la necesidad de recursos para resolverlo. Putnam parece concebir el lenguaje enteramente como un repositorio de teoría, y no en absoluto como una práctica de interacción social. Tenemos el lenguaje de la enciclopedia, pero ¿dónde está el lenguaje del pub? ¿Dónde están las intenciones comunicativas y las expectativas mutuas que parecen tener tanto que ver con lo que queremos decir? De hecho, ¿dónde está el pensamiento? Parece que entra en la imagen, si es que lo hace, solo como el caso especial donde el lenguaje a ser interpretado es innato (*hard-wired*), no hablado, escondido y demasiado conjetural.

Creo que el punto está bien formulado, pero considero que no tiene importancia. Si el problema del carácter de “pretendido” está bien propuesto, aún tendríamos la amenaza de la indeterminación radical, todavía habría necesidad de recurrir a restricciones salvadoras, aún habría un remedio análogo a la respuesta que Merrill sugiere para el planteamiento de Putnam y aún necesitaríamos propiedades naturales.

Déjese el lenguaje de lado y considérese en cambio la interpretación del pensamiento (más tarde podemos esperar comprender el lenguaje de un sujeto en términos de sus creencias y deseos con respecto a la comunicación verbal con los demás). El sujeto se encuentra en varios estados, y podría estar en varios más que están causalmente relacionados unos con otros, así como con el comportamiento del sujeto y con el entorno cercano que estimula sus sentidos. Estos estados encajan en una organización funcional, ocupan ciertos roles causales (muy probablemente son estados del cerebro. Quizá dicen relación con algo que es como el lenguaje, pero innato (*hard-wired*), o quizá no. Pero la naturaleza de los estados no interesa aquí). Estos tienen sus papeles funcionales en el sujeto tal como es ahora y en este tal como es en otros tiempos y tal como sería bajo otras circunstancias, e incluso en otras criaturas del mismo tipo que el sujeto. Dados los papeles funcionales de los estados, el problema es asignarles contenido (contenido proposicional, dirán algunos); pero yo estaría de acuerdo solo si las proposiciones pueden ser consideradas egocéntricas, y pienso que una “proposición egocéntrica” es simplemente una propiedad. Los estados indexados por contenido pueden ser identificados como una creencia cualquiera, como un deseo por aquello, como una experiencia perceptual de aparente confrontación

con tal-y-tal, como una intención de hacer tal-y-cual (pero no todas las adscripciones de actitudes ordinarias simplemente especifican el contenido de los estados del sujeto. Fred y Ted pueden ser iguales en relación con los papeles funcionales de sus estados y, en consecuencia, tener estados con el mismo contenido, en el sentido psicológico estrecho que es mi presente ocupación, y, por consiguiente, creer igualmente, *i.e.* al creer cada uno por sí mismo que ha escuchado sobre un bello pueblo llamado “Castlemaine”. Sin embargo, podrían estar familiarizados, a través de ese nombre, con pueblos distintos, ubicados en los extremos opuestos de la Tierra, de tal manera que Fred, y no Ted, cree que Castlemaine, Victoria, es bello).

El problema de asignar contenido a estados funcionalmente caracterizados debe resolverse por medio de principios restrictivos. Los principales entre estos son los principios de ajuste. Si un estado ha de ser interpretado como la intención de levantar la propia mano de uno, entonces típicamente ha de causar que la mano se levante. Si un estado (o un complejo de estados) ha de ser interpretado como un sistema de creencias y deseos –o, mejor aún, grados de creencias y deseos– de acuerdo con el cual levantar la mano sería un buen medio para los fines propios, y si otro estado ha de ser interpretado como una intención de levantar la mano, entonces el primero típicamente ha de causar el segundo. De la misma manera en el lado de la entrada (*input*). Un estado típicamente causado por una cosa redonda delante de los ojos es un buen candidato para interpretación como la experiencia visual de enfrentarse a algo redondo; y su impacto típico sobre los estados interpretados como sistemas de creencias debe ser comprendido como la adición exógena de la creencia de que uno está enfrentado a algo redondo, cualquiera que sea el ajuste que dicha adición requiera.

Hasta ahora, todo está bien, pero parece claro que malas interpretaciones absurdas y perversas pueden, no obstante, ser coherentes, esto es, podrían arreglárselas para ajustarse a los papeles funcionales de los estados, pues la asignación equivocada de contenido en un punto compensa la asignación equivocada en otro. Veamos cómo podría ocurrir esto, al menos bajo un esquema simplificado de interpretación como el que sigue. Una interpretación está dada por el par de funciones c y v . c es una distribución de probabilidad sobre los mundos, considerada como encapsulando las disposiciones del sujeto para formar creencias bajo el impacto de la evidencia sensorial: si una corriente de evidencia especificada por la proposición E pone a un sujeto en un estado total s –en corto, si E produce s –, nosotros interpretamos que S consiste en parte del sistema de creencias dado por la probabilidad de distribución c ($-/E$) que proviene de c al condicionar sobre E . v es una función que va desde los mundos a indicadores numéricos de preferibilidad, considerada

como encapsulando los valores básicos del sujeto: si *E* produce *s*, nosotros interpretamos que *S* consiste en parte en el sistema de deseos dados por las expectativas *C* ($-/E$) de *v*. Digamos que *c* y *v* *racionalizan* el comportamiento *B* después de la evidencia *E* si y solo si el sistema de deseos dado por las expectativas *C* ($-/E$) de *v* posicionan a *B* al menos tan alto como cualquier comportamiento alternativo. Digamos que *c* y *v* *se ajustan* si y solo si, para cualquier *E* que se encuentre especificando evidencia, *E* produce un estado que causaría un comportamiento racionalizado por *c* y *v* en virtud de *E*. Este es nuestro único principio restrictivo de ajuste. (¿A dónde fueron a parar los otros? –los hemos incorporado en las definiciones dondequiera que *c* y *v* encapsulen una asignación de contenidos de varios estados–). Entonces, dos interpretaciones cualesquiera que siempre racionalicen el mismo comportamiento en virtud de la misma evidencia deben ajustar igualmente bien. Llámese a dos mundos *equivalentes* si y solo si son iguales con respecto a la evidencia y comportamiento del sujeto, y nótese que cualquier mundo decente es equivalente *inter alia* a mundos horrendamente contrainductivos y a mundos en donde todo lo no observado por el sujeto es horrendamente desagradable. El ajuste depende del total de *c* para cada clase de equivalencia y de la expectativa-*c* de *v* dentro de cada clase, pero eso es todo. Dentro de una clase, no hace ninguna diferencia cuál mundo obtiene qué pares de *c* y *v*. Podemos intercambiar mundos equivalentes *ad lib* y preservar el ajuste. Por lo tanto, dada cualquier interpretación que se ajuste y sea razonable, podemos transformarla en una interpretación perversa igualmente ajustada al intercambiar los mundos equivalentes para mejorar las probabilidades de mundos contrainductivos, o la preferibilidad de mundos desagradables, o ambas. *Quod erat demonstrandum* (mis simplificaciones fueron horrendas: dejé afuera la egocentricidad de la creencia, el deseo y la evidencia, el aspecto causal del comportamiento racionalizado, el papel de las intenciones, el cambio de los valores básicos, las limitaciones de la competencia lógica, etc., pero dudo que estas omisiones tengan importancia para mi conclusión. Conjeturo que si fueran resueltas, todavía podríamos transformar interpretaciones razonables en perversas de una manera que permita preservar el ajuste).

Si confiamos en principios de ajuste para que hagan todo el trabajo, podemos esperar una radical indeterminación de la interpretación. Necesitamos restricciones adicionales, de aquellas como los llamados principios de caridad (sofisticada) o de “humanidad”.¹⁸ Dichos principios exigen interpretaciones según las cuales el sujeto tiene actitudes que estimaríamos razonables para uno que ha vivido la vida que ha vivido

18 Véase mi “Radical Interpretation” (1974) y el texto de Richard E. Grandy “Reference, Meaning and Belief” (1973).

(no como los principios de caridad cruda, que exigen imputaciones de error si él ha vivido bajo condiciones engañosas). Estos principios seleccionan entre interpretaciones en conflicto que se ajustan igualmente bien a los principios de ajuste. Ellos imponen presunciones *a priori* –aunque anulables– sobre qué clase de cosas son susceptibles de ser creídas o deseadas o, más bien, sobre qué disposiciones para desarrollar creencias y deseos, qué sesgos inductivos y valores básicos, se podrían correctamente interpretar.

Es aquí donde necesitamos propiedades naturales. Los principios de caridad impondrán un sesgo a creer que las cosas son verdes en vez de verdul, a desear una vida larga en vez de una vida-larga-a-menos-que-uno-haya-nacido-un-lunes-y-en-ese-caso-vida-por-un-número-par-de-semanas. En resumen, ellos imputarán contenido elegible, donde la inelegibilidad consiste en una severa no naturalidad de las propiedades que el sujeto supuestamente cree o desea o pretende tener para sí. Imputarán otras cosas también, pero es la elegibilidad imputada la que nos interesa en este momento.

Es así como la amenaza de la indeterminación radical en la asignación de contenido del pensamiento es repelida. La restricción salvadora indica una relación con el contenido –no con el pensador ni con cualquiera de los canales entre ambos–. Se requieren dos para indexar estados con contenido, y no encontraremos la restricción si seguimos buscando en el lado equivocado de la relación. Creer esto o desear aquello consiste en parte en los papeles funcionales de los estados por los cuales creemos o deseamos, pero en parte consiste en la elegibilidad del contenido. Y esta elegibilidad a ser considerada es un asunto, en parte, de propiedades naturales.

Considérese el *puzzle* donde Kripke ilustra la paradoja de Wittgenstein de que “ningún curso de acción puede ser determinado por una regla, porque cada curso de acción puede hacerse ajustado a la regla”.¹⁹ Una persona bien educada que trabaja problemas de aritmética pretende realizar una suma cuando ve el signo “+”. Él no pretende llevar a cabo una cuadición (*quaddition*), que es una clase de suma para números pequeños, pero que da como resultado el número 5, si cualquiera de los números sometidos a la cuadición excede cierto límite. ¿Por qué él pretende sumar y no cuadicionar? Lo que sea que diga y lo que sea que escriba en su cerebro puede ser perversamente (mal) interpretado como dándole instrucciones para que cuadicione. Y no es suficiente decir que su estado cerebral es la base causal de la disposición para sumar. Quizá no lo es. Tal vez si surgiera un caso de prueba él abandonaría su

19 Véase Saul A. Kripke (1981)

intención, no sumaría ni cuadicionaría, sino más bien pondría su tarea de lado y alegraría que los problemas son muy difíciles.

La solución ingenua consiste en decir que sumar significa seguir de la misma manera que antes a medida que los números se van haciendo más grandes, en cambio cuadicionar significa hacer algo diferente; no hay nada presente en el sujeto que constituya una intención de hacer cosas distintas en diferentes casos; por lo tanto, él pretende hacer la suma, no la cuadición. No deberíamos burlarnos de esta respuesta ingenua. Es la solución correcta a este *puzzle*. Pero debemos pagar por recuperar nuestra ingenuidad. Nuestra teoría de las propiedades debe tener los recursos adecuados para ratificar de alguna manera el juicio de que las instancias de las sumas son todas iguales de una manera en que las instancias de la cuadición no lo son. La propiedad de sumar no es perfectamente natural, por supuesto, no a la par con la unidad de carga o la esfericidad. Y la propiedad de cuadicionar no es perfectamente no-natural. Pero cuadicionar es peor por una disyunción. Por lo tanto, la cuadición es, en ese respecto, menos un hacer lo mismo y, debido a ello, es en ese respecto una cosa menos elegible como algo que uno tendría la intención de hacer.

No es que usted posiblemente no pueda tener la intención de cuadicionar. Usted podría hacerlos. Supóngase que hoy hay tanta base como puede haber para interpretarlo a usted como pretendiendo sumar y como queriendo decir adición con la palabra “adición” y cuadición con la palabra “cuadición”; y mañana se dice a sí mismo claramente que sería divertido fastidiar a los filósofos solo tomando la cuadición de aquí en adelante, y se decide a hacerlo. Pero será necesario salirse del camino. Sumar y cuadicionar no están en el mismo nivel. Para pretender sumar, solo es necesario tener estados que se ajusten a cualquiera de las interpretaciones y dejarle a la caridad que decreta que usted tiene intención más elegible. Para pretender cuadicionar, usted debe decir o pensar algo que crea dificultades de ajuste para la intención más elegible y, por lo tanto, destruye la presunción a su favor. Debe hacer algo que, tomando conjuntamente los principios de ajuste y las presunciones de elegibilidad y otros principios de caridad, incline la balanza a favor de que usted pretende cuadicionar. ¡Qué irónico que hayamos estado preocupados por no encontrar nada positivo que resolviera el asunto a favor de la adición! Es la falta de algo positivo que apunte en cualquier dirección *lo* que apunte a favor de la adición. La cuadición, siendo menos natural y elegible, requiere algo positivo a su favor. La adición siempre puede ganar por defecto.

¿Cuál es el estatuto de los principios que restringen la interpretación, en particular la presunción caritativa a favor del contenido elegible? Debemos rehuir varios malos entendidos. No debe decirse (1) que,

como hecho psicológico contingente, los contenidos de nuestros estados resultan ser bastante elegibles, nosotros en la mayoría de los casos creemos y deseamos tener propiedades no tan no-naturales. Menos aún deberíamos decir (2) que nosotros deberíamos atrevidamente presuponer esto en nuestra interpretación de unos a otros, incluso si ni siquiera tenemos un poco de evidencia para ello. Tampoco debería ser dicho (3) que, como hecho psicológico contingente, resulta que tenemos estados cuyo contenido involucra algunas propiedades en lugar de otras, y eso hace que las primeras propiedades sean más naturales (esto sería una teoría psicologista de la naturalidad). El error es el mismo en los tres casos. Se supone, incorrectamente a mi juicio, que el problema de la interpretación puede ser resuelto sin traer a colación la distinción entre propiedades naturales y propiedades no naturales; de tal manera que las propiedades naturales pueden o no resultar como aquellas presentes en el contenido de nuestro pensamiento de acuerdo con la solución correcta, o de tal manera que después puedan ser definidas como aquellas que figuran de esa manera. Creo que esto es demasiado optimista. No tenemos idea sobre cómo resolver el problema de la interpretación mientras consideremos todas las propiedades como igualmente elegibles para ser contenido del pensamiento. Eso implicaría resolver el problema sin tener restricciones suficientes. Solo si tenemos una distinción independiente, objetiva entre propiedades, e imponemos la presunción a favor de un contenido elegible *a priori* como una restricción constitutiva, el problema de la interpretación tiene una solución. Si esto es así, entonces cualquier solución debe automáticamente respetar la presunción. Aquí no hay ningún dato contingente de la psicología que creer, ya sea con evidencia o atrevidamente.

Compárese nuestra teoría selectiva y colectiva de la legalidad: la legalidad de una regularidad consiste en su ajuste a un sistema ideal de alta puntuación, por lo que es inevitable que las leyes resulten tener lo necesario para producir altas puntuaciones. Así mismo, como he sugerido, el carácter de ser contenido consiste en ser asignado por una interpretación de alta puntuación, por lo que es inevitable que los contenidos tiendan a tener lo necesario para producir puntuaciones altas. En ambos casos, he sugerido que parte de lo que se necesita es la naturalidad de las propiedades involucradas. La razón por la cual las propiedades naturales figuran en los contenidos de nuestras actitudes, es que la naturalidad es parte de lo que es figurar en ellas; no es que estemos contruidos para tener un interés especial en las propiedades naturales, o que nosotros le otorguemos la naturalidad a las propiedades cuando tomamos interés en ellas.²⁰

Bibliografía

- Armstrong, D. M. *Universals and Scientific Realism*. 11 vols. Cambridge: Cambridge University Press, 1978.
- Armstrong, D. M. *What is a Law of Nature?* Cambridge: Cambridge University Press, 1983. <http://dx.doi.org/10.1017/CBO9781139171700>.
- Bennett, J. "Review of *Counterfactuals*." *Canadian Journal of Philosophy* 4 (1974): 381-402.
- Downing, P. B. "Subjunctive Conditionals, Time Order and Causation." *Proceedings of the Aristotelian Society* 59 (1959): 125-140.
- Dretske, F. "Laws of Nature." *Philosophy of Science* 44 (1977): 248-268.
- Gibbard, A., and Harper, W. "Counterfactuals and Two Kind of Expected Utility." *Foundations and Applications of Decision Theory*. Eds. Clifford Hooker, James Leach and Edward McClennen. Dordrecht; Boston: Reidel, 1978. 341-376.
- Grandy, R. "Reference, Meaning and Belief." *Journal of Philosophy* 70 (1973): 439-452.
- Horgan, T. "Supervenience and Microphysics." *Pacific Philosophical Quarterly* 63 (1982): 29-43.
- Inwagen, P. "The Incompatibility of Free Will and Determinism." *Philosophical Studies* 27 (1975): 185-199.
- Jackson, F. "A Causal Theory of Counterfactuals." *Australasian Journal of Philosophy* 55 (1977): 3-21.
- Kim, J. "Supervenience and Nomological Incommensurables." *American Philosophical Quarterly* 15 (1978): 149-156.
- Kripke, S. "Identity through Time." *Diss. American Philosophical Association, Eastern Division*, 1979.
- Kripke, S. "Wittgenstein on Rules and Private Language: An Elementary Exposition." *Perspectives on Wittgenstein*. Ed. Irvin Block. Oxford: Basil Blackwell, 1981. 238-312.
- Lewis, D. "Extrinsic Properties." *Philosophical Studies* 44 (1983): 197-200. <http://dx.doi.org/10.1007/BF00354100>.
- Lewis, D. "Counterfactual Dependence and Time's Arrow." *Noûs* 13 (1979): 455-476.
- Lewis, D. *Counterfactuals*. Oxford: Blackwell, 1973.
- Lewis, D. *Causation and Conditionals*. Ed. Ernest Sosa. Oxford University Press, 1975. Rpt. of "Causation." *Journal of Philosophy* 70 (1973): 556-567.
- Lewis, D. "Radical Interpretation." *Synthese* 23 (1974): 331-334.

20 Estoy agradecido por los comentarios de Gilbert Harman, Lloyd Humberstone, Frank Jackson, Mark Johnston, Donald Morrison, Kim Sterelny, y otros; y especialmente a la discusión y correspondencia con D. M. Armstrong en los últimos años, sin la cual seguiría creyendo hasta este día que la teoría de conjuntos aplicada a los possibilia es toda la teoría de propiedades que uno puede necesitar.

- Merrill, G. "The Model-Theoretic Argument against Realism." *Philosophy of Science* 47 (1980): 68-81. <http://dx.doi.org/10.1086/288910>.
- Montague, R. "Deterministic Theories." *Decisions, Values and Groups* II. Oxford: Pergamon Press, 1962. 325-370.
- Montague, R. *Formal Philosophy: Selected Papers of Richard Montague*. New Haven: Yale University Press, 1974.
- Prior, E., Pargetter, R., and Jackson, F. "Three Theses about Dispositions." *American Philosophical Quarterly* 19 (1982): 251-257.
- Putnam, H. "Realism and Reason." *Meaning and the Moral Sciences*. London: Routledge and Kegan Paul, 1978. 123-140
- Putnam, H. "Models of Reality." *The Journal of Symbolic Logic* 3.45 (1980): 464-482. <http://dx.doi.org/10.2307/2273415>.
- Ramsey, F.P. "Universals of Law and of Fact." *Foundations*. London: Routledge and Kegan Paul, 1978. 128-132
- Schiffer, S. "The Basis of Reference." *Erkenntnis* 13 (1978): 171-206.
- Slote, M. "Some Thoughts on Goodman's Riddle." *Analysis* 4.27 (1967): 128-132.
- Stole, M. *Reason and Scepticism*. London: George Allen and Unwin, 1970.
- Stuart Mill, J. *A System of Logic*. London: Parker, 1843
- Tooley, M. "The Nature of Laws." *Canadian Journal of Philosophy* 4 (1977): 667-698.