



<http://dx.doi.org/10.15446/ideasyvalores.v66n163.49743>

# ¿SOLUCIONA EL DESCRIPTIVISMO CAUSAL EL PROBLEMA DE LA REFERENCIA DE LOS TÉRMINOS TEÓRICOS?\*



## DOES CAUSAL DESCRIPTIVISM SOLVE THE PROBLEM OF REFERENCE OF THEORETICAL TERMS?

BRUNO BORGE\*\*

Universidad de Buenos Aires / ANPCYT - Buenos Aires - Argentina

.....  
*Artículo recibido el 20 de marzo del 2015; aprobado el 13 de mayo del 2015.*

\* El presente trabajo fue realizado con el apoyo de la Agencia Nacional de Promoción Científica y Tecnológica de Argentina (ANPCYT). Algunos de los argumentos correspondientes a la quinta sección fueron presentados en el marco del III Coloquio de Filosofía e Historia de la Ciencia, celebrado en 2015 en la ciudad de Montevideo, Uruguay. Agradezco al auditorio por sus valiosas observaciones críticas. En el mismo sentido, quisiera expresar mi agradecimiento hacia un evaluador anónimo de esta revista, cuyos comentarios permitieron mejorar la versión final de este trabajo.

\*\* [brunojborge@gmail.com](mailto:brunojborge@gmail.com)

### Cómo citar este artículo:

MLA: Borge, B. "¿Soluciona el descriptivismo causal el problema de la referencia de los términos teóricos?" *Ideas y Valores* 66.163 (2017): 125-151.

APA: Borge, B. (2017). ¿Soluciona el descriptivismo causal el problema de la referencia de los términos teóricos? *Ideas y Valores*, 66 (163), 125-151.

CHICAGO: Bruno Borge. "¿Soluciona el descriptivismo causal el problema de la referencia de los términos teóricos?" *Ideas y Valores* 66, n.º 163 (2017): 125-151.



This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.

### RESUMEN

Las teorías de la referencia puramente descriptivistas o causales fracasan a la hora de dar cuenta del modo en que se fija y puede rastrearse la referencia de los términos teóricos. Psillos propuso dos versiones del descriptivismo causal que recogen argumentos presentes en defensas previas de dicha posición. Se trata de una teoría mixta que pretende solucionar el problema y acomodarse a intuiciones presentes en enfoques alternativos, como el que apela a oraciones de Ramsey. El artículo se propone mostrar la insuficiencia del descriptivismo causal como solución al problema de la referencia de los términos teóricos.

*Palabras clave:* S. Psillos, descriptivismo causal, oración de Ramsey, realismo científico, términos teóricos.

### ABSTRACT

Purely descriptivist or purely causal reference theories fail to account for the manner in which the reference of theoretical terms is fixed and can be traced. S. Psillos proposed two versions of causal descriptivism that take up arguments set forth in previous defenses of said position. His is a mixed theory that attempts to solve the problem by adopting insights from alternative approaches, such as those that resort to Ramsey sentences. The objective of the article is to show the insufficiency of causal descriptivism as a solution to the problem of reference of theoretical terms.

*Keywords:* S. Psillos, causal descriptivism, Ramsey sentences, scientific realism, theoretical terms.

## Introducción

El problema de la referencia de los términos teóricos constituye sin dudas uno de los tópicos centrales de la filosofía de las ciencias. Varias son las razones que le han valido dicha centralidad. En primer lugar, la posición que se adopte al respecto define en buena medida el bando que se defiende en la polémica realismo vs. antirrealismo científico, puesto que el realismo científico suele definirse como la tesis que sostiene que los términos teóricos de nuestras mejores teorías científicas refieren.<sup>1</sup> Además, la actitud realista hacia las teorías científicas está profundamente ligada a la afirmación de una continuidad epistémica a lo largo del cambio teórico. En otras palabras, al hecho de que al menos algunas piezas del vocabulario de teorías abandonadas referían exitosamente a entidades inobservables (precisamente, aquellas entidades sobre las que las nuevas teorías han aumentado nuestro conocimiento). Los antirrealistas, por su parte, no han escapado al problema de construir un relato sobre el significado de los términos teóricos. Desde los monumentales esfuerzos de los miembros del Círculo de Viena por reducirlos a un vocabulario puramente observacional, hasta los múltiples intentos de Kuhn por precisar su tesis de la inconmensurabilidad semántica, toda una batería de argumentos se ha ofrecido para ilustrar plausiblemente por qué la pretendida continuidad referencial de los realistas es un objetivo inalcanzable. Por otra parte, la disputa tradicional en torno a la interpretación de las teorías científicas se vio alimentada por una variedad de posiciones encontradas acerca de cuestiones como la denotación y el significado. En particular, por la irrupción de la teoría causal de la referencia de Kripke (1980) como una alternativa a las propuestas descriptivistas que constituían, en mayor o menor medida, variaciones de la tradición fregeana de análisis del fenómeno del significado. La posibilidad de que la referencia de los términos teóricos estuviese determinada de manera directa por la entidad que designan, resultó sumamente atractiva para quienes pretendían desarticular argumentos discontinuistas como la tesis del holismo semántico. El resultado, sin embargo, fue que las nuevas armas no resultaron más efectivas contra los argumentos antirrealistas que las clásicas soluciones descriptivistas. El hecho de que el bautismo de una magnitud o entidad inobservable

- .....
- 1 Si bien esta no es la única manera de presentar el realismo científico, ha sido una de las más frecuentes en los inicios del debate, cuando las nociones propias de la concepción sintáctica eran moneda corriente. La concepción semántica de las teorías –para la cual una teoría científica es una familia de modelos, y no un conjunto de proposiciones– ha procurado desalojar dichas nociones de las consideraciones epistemológicas, sin embargo, se ha argumentado plausiblemente que el problema de la referencia perdura bajo una nueva forma en el marco de la concepción semántica. Véase, por ejemplo, Halvorson (2012) y Lutz (2014).

no pueda realizarse ostensivamente, implica que el término que la designa acaba por referir a *lo que fuere que produzca* los fenómenos que se pretenden explicar. Mientras que las teorías descriptivistas relativizan el significado de los términos teóricos al contexto de la teoría en la que se los utiliza, las alternativas causales terminan por hacer del éxito referencial una cuestión trivial. En ese marco muchos han apostado por una posición que pretende reunir lo mejor de ambas alternativas, una teoría híbrida de la referencia denominada *descriptivismo causal*. En su primera versión, este sostiene que la referencia de los términos teóricos está determinada, por una parte, por un componente causal que vincula los fenómenos con la entidad postulada para dar cuenta de dichas observaciones, pero, por otra parte, por un subconjunto de las descripciones teóricas asociadas a dicha entidad, a saber, aquellas que consignan las propiedades constitutivas de clase por las que la entidad inobservable juega el rol causal que la teoría le adjudica. En otra variante de esta misma posición, las propiedades causales que determinan la referencia no tienen que coincidir con las propiedades constitutivas de clase de la entidad postulada. Psillos ha defendido ambas variantes a lo largo de su obra; asimismo, se ha ocupado de mostrar los puntos de contacto entre el descriptivismo causal y las alternativas que han intentado abordar el problema de la referencia de los términos teóricos desde el marco que otorga el análisis de Ramsey del contenido teórico (cf. 1999; 2012b).

El descriptivismo causal, según sus defensores, comporta varias ventajas sobre otras teorías a la hora de comprender el modo en que adquieren y mantienen su referencia los términos teóricos. En el presente trabajo me propongo poner en cuestión este punto, mostrando que al menos dos de sus versiones más acabadas fracasan en cuanto teorías de la referencia para los términos teóricos, incapaces de satisfacer los *desiderata* del realismo científico. En la segunda sección expongo de manera general las razones por las cuales las teorías de la referencia puramente descriptivistas o puramente causales no resultan adecuadas en el caso específico de los términos teóricos. En la tercera sección presento una caracterización del descriptivismo causal en general y de su aplicación dentro de la filosofía de la ciencia. En ese contexto, abordo la primera de sus variantes con el propósito de mostrar la imposibilidad de establecer algunos de sus supuestos básicos. En la cuarta sección me ocupo de una versión alternativa de esta teoría, tal como aparece en la segunda etapa de la obra de Psillos. Si bien esta nueva propuesta subsana algunas falencias presentes en la primera, muestro que no consigue dar cuenta de cómo se fija y rastrea la referencia de los términos teóricos. En la quinta sección analizo una estrategia alternativa respecto de dicho problema, fundada en la posibilidad de eliminar los términos teóricos

por medio de oraciones de Ramsey. El objetivo principal de este análisis será cuestionar su compatibilidad con los principios del descriptivismo causal, lo que refuta las afirmaciones de Psillos al respecto. No obstante, el curso de mi argumentación permite exponer algunos puntos de contacto entre la alternativa de Ramsey, la objeción de Newman al realismo estructural y el argumento modelo-teórico contra el realismo metafísico. Por último, la sexta sección incluye las conclusiones y comentarios finales a este trabajo.

### Teorías de la referencia y realismo científico

La teoría descriptivista de la referencia sostiene que una expresión referencial (por ejemplo un nombre propio o un término singular) obtiene su referencia a partir de la mediación de una descripción (o un *cúmulo* de descripciones), normalmente entendida como especificación del *sentido* de dicha expresión. Así, si una cierta entidad  $x$  satisface la descripción (o el *cúmulo* de descripciones, o las descripciones relevantes de ese *cúmulo*) asociada a un cierto término  $t$ , entonces  $t$  refiere a  $x$ ; si en cambio ninguna entidad satisface esa descripción (o el *cúmulo* de descripciones, o las descripciones relevantes de ese *cúmulo*) asociada a  $t$ , entonces  $t$  no refiere. Debe notarse que esta no es solo una teoría acerca del significado de términos como  $t$ , sino también acerca del modo en que su referencia queda fijada. Aunque no exentas de problemas, variantes de esta teoría respondieron de manera satisfactoria a algunos de los desafíos planteados por Frege, y en parte por ello dominaron el panorama por décadas, hasta la aparición de la teoría causal de la referencia en la escena filosófica.

No es difícil ver en qué sentido este relato resulta altamente problemático para el realista científico. Como sugerí anteriormente, una de las metas propias del realismo es poder brindar una reconstrucción del fenómeno de la referencia de los términos teóricos que garantice una cierta continuidad referencial a lo largo del cambio teórico, dando cuenta del carácter acumulativo del conocimiento científico. El descriptivismo no parece poder cumplir dichas expectativas: un cambio de teoría, aun cuando el uso de ciertos términos se preserve, conlleva un cambio sustancial en muchas de las descripciones asociadas a los términos teóricos que se incluían en la teoría abandonada, y tenidas esas descripciones por falsas, no hay modo en el que la continuidad referencial pueda sustentarse. Si bien esto puede ser plausible en el caso de *algunas* entidades postuladas por teorías pretéritas, no parece ajustarse en general a buena parte de las pretensiones realistas. Las descripciones que Dalton asociaba al término “átomo” eran sustancialmente diferentes a las que la teoría atómica contemporánea aceptaría suscribir; por tanto, cuando Dalton hablaba de “átomos” no refería a los átomos,

sino más bien a nada. Estas implicaciones fueron recogidas por Kuhn a la hora de sentar las bases de la tesis de la inconmensurabilidad entre teorías (cf. 1962). A la variación del significado propia del cambio de descripciones asociadas a los términos teóricos, se sumó la afirmación de un holismo semántico que recorría todas las dimensiones del lenguaje científico; así, entre teorías o marcos conceptuales diferentes existe una variación tan radical de significado que no existe base común para su comparación. Esta inconmensurabilidad semántica tiene su correlato ontológico en un cambio en la referencia del lenguaje científico como totalidad. En gran medida por ello Kuhn sentencia que, “cuando los paradigmas cambian, el mundo mismo cambia con ellos” (1962 111).<sup>2</sup> El papel central del holismo semántico en el planteo radical de Kuhn puede apreciarse más claramente si se compara su posición con otras propuestas antirrealistas. El fallido proyecto del positivismo lógico de brindar definiciones explícitas de los términos teóricos en un vocabulario puramente observacional encontró una forma más moderada en las “oraciones reductivas” de Carnap (1936), y aunque en ambos casos la teoría de la referencia de base es descriptivista, las variaciones del significado de los términos teóricos tienen un punto arquimédico en la neutralidad del vocabulario observacional. Es por ello que el holismo semántico tiene implicaciones tan profundas, pues, luego de un cambio de teoría, todos los términos del lenguaje cambian de significado, incluso aquellos que refieren a la experiencia inmediata.

La teoría causal de la referencia presentada por Kripke (1980), por su parte, sostiene que una expresión referencial puede referir exitosamente a un individuo, aun cuando muchas de las descripciones (incluso *todas* las descripciones) que los hablantes asocian a dicha expresión resulten ser falsas; del mismo modo que, inversamente, un individuo puede satisfacer una clase de descripciones asociadas a un nombre y aun así no ser el referente de ese nombre. El relato positivo acerca de cómo un término adquiere su referencia sería algo similar a lo siguiente: en un “acto de bautismo”, un hablante *etiqueta* una cierta entidad con un nombre, esa conexión mediante la cual la referencia queda fijada y se transmitirá al resto de los hablantes por medio de una cadena causal que remitirá (trazando el camino inverso) a ese acto de bautismo inicial. Puede que muchas o incluso todas las descripciones que los hablantes asocien al nombre sean erradas, pero en cualquier caso eso no altera el vínculo original entre un nombre y su referente, en cuanto este es un vínculo directo, no mediado por descripción o sentido alguno. Kripke admite incluso la posibilidad de que el acto de fijación de la referencia sea efectuado *mediante* una descripción, como en el caso del bautismo

.....  
2 A menos que se indique lo contrario, todas las traducciones son de mi autoría.

del planeta Neptuno, pero aun así esa descripción agota su rol en el de un instrumento que permite establecer el vínculo (cf. 1980 96). El punto crucial es, nuevamente, que la descripción tiene una relación contingente con el referente.

Putnam hizo notar que la referencia de los términos que designan clases naturales o magnitudes físicas puede ser explicada de modo análogo (cf. 1975; 1983). Dada una muestra de cierta sustancia, puede nombrarse por ostensión y fijar la referencia para todas las muestras posibles de esa sustancia, es decir, para todas las instancias de esa clase. La referencia se transmite entonces de hablante en hablante mediante una cadena causal, que remite al acto de bautismo por el cual el nombre queda fijado para toda muestra o ejemplar que comparta la misma naturaleza con la muestra o ejemplar presente en ese acto. En palabras de Putnam: “un término refiere (al objeto nombrado) si se encuentra en la relación correcta (continuidad causal en el caso de los nombres propios; identidad de ‘naturaleza’ en el caso de los términos de clase) con aquellas cosas existencialmente dadas” (1983 73). Una estrategia semejante, cabe agregar, podría seguirse para dar cuenta de la referencia de los términos teóricos: ante la presencia de ciertos fenómenos (e. g. una estela fotografiada en una cámara de burbujas) el científico se ve en la necesidad de postular alguna entidad inobservable que sea causalmente responsable de ellos (e. g. un electrón). Obviamente, el acto de bautismo no involucra aquí una fijación ostensiva de la referencia, sino mediada por una o varias descripciones, en principio, la que señala que es la entidad causante de esos fenómenos.

Las consideraciones anteriores parecen muy promisorias para el realista científico. La razón de ello radica en que recogen una intuición que le es bastante conveniente, *i. e.* que uno *puede* referirse a cierta entidad sin conocer cuál sería una descripción completa o suficientemente adecuada de ella, y aun cuando algunas (o muchas) de las cosas que crea acerca de esa entidad estén erradas. El resultado de la investigación empírica acerca de una entidad inicialmente postulada como principio explicativo de la ocurrencia de ciertos fenómenos, seguramente será una serie de modificaciones de las descripciones teóricas originalmente ensayadas, pero eso no nos impide suponer que se trata de descripciones de *la misma* entidad.

Sin embargo, este relato presenta algunos problemas insalvables. En el caso de la referencia de los términos teóricos, la ostensión –a diferencia de lo que ocurre respecto de los términos de clases naturales manifiestas– no cumple ningún papel en la fijación de la referencia. Esta, por el contrario, se realiza únicamente mediante una descripción teórica. Dado un fenómeno (o conjunto de fenómenos), se identifica a este como el efecto de alguna causa subyacente (una cierta entidad o magnitud física) que recibe su

nombre en ese acto de bautismo efectuado mediante la descripción causal correspondiente. Según el esquema de Putnam, todas las ocurrencias de un mismo término teórico (incluso en teorías diferentes) remiten por una cierta cadena causal a la circunstancia en que la descripción causal de la entidad fue dada, asignando a los fenómenos observables como sus efectos. La entidad postulada no puede serlo sino como *aquel agente causal* responsable de los efectos observados, es decir, su postulación es meramente existencial. El problema salta inmediatamente a la vista: una teoría de la referencia directa se vería trivialmente satisfecha para cualquier fenómeno bajo la sola asunción de que este tiene una causa. Dicho de otro modo, en el marco de esta teoría no es posible dar cuenta de cómo la referencia de un término teórico podría fallar. Y este resultado es, aun para el realista más ortodoxo, poco satisfactorio. El término “flogisto”, por ejemplo, fue introducido para nombrar aquella entidad que cumple de hecho el mismo rol causal que el oxígeno, pero sabemos que el flogisto no existe. Pero si aceptamos el relato causal, no podemos menos que concluir cosas tales como que “flogisto”, en lugar de no denotar, refiere al oxígeno, o que “acción gravitacional a distancia” refiere a la curvatura espacio-temporal. Y en el caso de querer escapar de esas conclusiones, parece que no tenemos otra opción que admitir que al menos *algunas* descripciones teóricas están involucradas no solo en la fijación de la referencia de los términos teóricos. Esta objeción –sostenida en repetidas ocasiones por filósofos de la ciencia (Bird 2000 185; Kroon 1985; Niiniluoto 1999, 126; Nola 1980; Psillos 1999 280)–, fue formulada por primera vez por Encé:

La tesis Kripke-Putnam [...] nos compele a afirmar que los teóricos del flogisto estaban de hecho hablando acerca del oxígeno, y que tenían algunas creencias falsas acerca del oxígeno. [...] Pero creo que esta conclusión es totalmente inaceptable. Me parece suficientemente claro que Lavoisier, al descubrir el oxígeno, descubrió que el flogisto no existe. (1976 268)

De ese modo, la teoría causal de la referencia trivializa el éxito referencial, y a la hora de adoptarla como una herramienta para abordar el problema de la denotación de los términos teóricos, el realista se topa con la indeseable consecuencia de no poder dar cuenta de los fallos de referencia en la historia de la ciencia.

Las reconstrucciones de las dos principales teorías de la referencia que he ensayado en esta sección han omitido sin duda buena parte de las complejidades y matices de sus variantes, sin embargo, exponen las limitaciones inherentes a toda propuesta puramente descriptivista o causal a la hora de brindar un relato acerca de la referencia de los términos teóricos acorde a las intuiciones realistas. El reconocimiento de dichas limitaciones ha sido clave para impulsar el desarrollo de teorías híbridas de la referencia.



### Descriptivismo causal, acto I: propiedades esenciales

Como ya he señalado, el mismo Kripke admite que al menos en ciertos casos la referencia de un término puede ser fijada mediante una descripción, sin embargo, dado que su papel en ese caso se limita al de un instrumento para establecer el vínculo causal entre el nombre y su referencia, no hace falta que dicha descripción sea verdadera ni que se transmita exitosamente para cumplir su propósito. La descripción “la estrella matutina” pudo ser usada para fijar la referencia de “Phosphorus”, aun cuando Phosphorus no sea una estrella (cf. Kripke 1980 79). Esta es una consecuencia de poner todo el peso de la referencia en la relación causal entre un término y su referente, pero esa estrategia tiene, como hemos visto, consecuencias indeseables a la hora de dar cuenta de la referencia de los términos teóricos. Una teoría que revitalice el rol de las descripciones en la fijación de la referencia, sin soslayar la importancia de los componentes causales, debe admitir que, al menos hasta cierto punto, las descripciones que se utilizan para fijar la referencia han de ser correctas. Los defensores de esta idea sostienen una posición denominada *descriptivismo causal*, debido a Lewis (1984). Por su parte, Loar (1976), Kroon (1987; 2008) y Jackson (1998) defendieron propuestas semejantes como alternativas a la teoría causal pura. La idea básica de esta posición es que los hablantes asocian a un nombre “N” una descripción de la forma DC, que afirma que existe una entidad que está en la relación R con mi uso actual del nombre “N”, donde R es construida de acuerdo al relato estándar de la transmisión de la referencia según la teoría causal. De ese modo, DC determina la referencia de “N”. Esta teoría, sin embargo, no resulta de mucha utilidad para el problema de los términos teóricos. Si dichos términos son introducidos por una descripción que postula solo existencialmente un agente responsable de ciertos fenómenos, poco importa si los hablantes hacen o no uso de una descripción del tipo DC. En suma, si ejemplos como el de “flogisto” son un problema para la teoría causal, también lo son para el descriptivismo causal. La inconveniencia de esta posición (es decir, del descriptivismo causal entendido como una teoría de la referencia para nombres propios) para el problema aquí considerado se hace patente, además, en algunas de las inflexiones de sus formulaciones particulares. Kroon subraya que la introducción de nombres y nuestras prácticas asociadas a ellos requieren de la *disponibilidad cognitiva* de los referentes (cf. 2008 146). Además, en el caso de las clases naturales, los nombres denotan “en virtud de que [el referente] resulta cognitivamente disponible para los usuarios por medio de la percepción” (Kroon 2008 146). No es difícil ver cómo esta pretensión de un acceso epistémico a los referentes complica la extensión de esta posición al caso de los términos teóricos.

Estas dificultades decantaron en la formulación de una versión específica del descriptivismo causal para los términos teóricos, que llamaré *descriptivismo causal-T* (de aquí en adelante, DCT).<sup>3</sup> Enc (1976), Nola (1980), Sankey (1997), Bird (2000) y Psillos (1999; 2012b) han defendido variantes de esta posición. Sus lineamientos podrían dibujarse del siguiente modo: como ya se ha dicho, la referencia de un término teórico  $t$  no puede ser fijada ostensivamente, sino por medio de descripciones teóricas. Dado que es la aparición de ciertos fenómenos  $\Phi$  lo que hace necesaria la postulación de la entidad denotada por  $t$ , parte de una descripción de  $t$  consistirá en la afirmación de que ella es causalmente responsable de la ocurrencia de  $\Phi$ . Pero, como también se ha hecho notar, eso resulta insuficiente. Lo que ha de fijar la referencia de un término teórico no será, sin embargo, una descripción completa de la entidad postulada, sino descripciones causales que den cuenta del modo efectivo en que la entidad cumple su rol causal respecto de  $\Phi$ .

La idea *negativa* detrás de esta teoría es que el mecanismo de fijación de la referencia no puede estar limitado a una relación causal no conceptual entre un término teórico y su referente, pero tampoco puede tratarse de una relación puramente conceptual basada en descripciones. La idea *positiva* es que el mecanismo de fijación de la referencia ha de tener la siguiente forma:

$R(x) = x$  causa los fenómenos  $\Phi$  y  $D(x)$ .

El término  $t$  refiere a  $x$  si y solo si  $R(x)$ .

El componente causal está dado por  $\neg x$  causa  $\Phi$ , es decir, por el hecho de que se pretende que el término teórico designe una entidad que se postula como la causa de algunos fenómenos. El componente descriptivo corresponde a  $D(x)$ , una descripción causal que explicita el modo en que el referente postulado se supone causalmente conectado con el fenómeno  $\Phi$ .

El premio por devolverle a las descripciones su perdido protagonismo es poder reconstruir un relato racionalmente aceptable acerca de la continuidad y discontinuidad en la referencia de algunos términos. Por ejemplo, en la  $D(x)$  correspondiente a “flogisto” ha de incluirse no solo su carácter de condición necesaria para la combustión, sino el hecho de que es “la sustancia liberada en la combustión”, y esa especificación de su rol causal asegura que nunca pueda compartir referente con “oxígeno”, en cuya  $D(x)$  no se incluye dicha descripción.

3 Esta distinción no está presente en la literatura especializada, en donde el rótulo de “descriptivismo causal” es aplicado indistintamente, tanto en el contexto de la referencia de los nombres propios o clases manifiestas, como en el caso de los términos teóricos. Espero que las consideraciones hasta aquí vertidas y los desarrollos de las secciones subsiguientes se juzguen suficientes para mostrar la pertinencia de trazar tal distinción.

Esta caracterización general del DCT admite dos importantes distinciones. La primera de ellas no concierne al modo en que la posición es formulada, sino a cómo dicha formulación es interpretada. Para un primer grupo, el DCT es una teoría híbrida, su componente descriptivo está dado por el hecho de que la referencia es fijada y determinada por una descripción, pero esa descripción tiene un carácter tal que aporta además un elemento causal. Para un segundo grupo, el DCT no debe ser interpretado como una alternativa genuina a las dos teorías en pugna. En esa línea, Bird afirma que el DCT en realidad “es una teoría puramente descriptivista, cuya característica es que algunas descripciones deben establecer que la entidad a la que se pretende referir juega un rol causal respecto de algunos fenómenos específicos” (2000 186). Una lectura semejante es ofrecida por Raatikainen (2007). Dado que el foco de este trabajo está puesto sobre la interpretación correspondiente al primer grupo, solo me detendré en dos breves comentarios. En primer lugar, los problemas del puro descriptivismo no pueden solucionarse con la adición de más descripciones. Como Sankey ha hecho notar, solo una versión de la teoría causal puede garantizar la estabilidad de la referencia ante el cambio conceptual (*cf.* 1997 428). En segundo término –y más importante–, es posible caracterizar al DCT como una modificación de la teoría causal si esta es adecuadamente entendida. Para Raatikainen el descriptivismo causal colapsa en el descriptivismo a secas, puesto que el único sentido en el que la *nueva teoría de la referencia* (como llama al programa inaugurado por Kripke y Putnam) es “causal” refiere a la transmisión de la referencia entre los hablantes, y no al modo en que esta queda fijada (*cf.* 2007 7). Discrepo con esta imagen. La fijación de la referencia es también causal en un sentido *negativo* y en uno *positivo*. *Negativo* en cuanto no requiere de la mediación necesaria de ninguna descripción; *positivo*, dado que las cadenas causales de transmisión de la referencia no comienzan necesariamente con la introducción de un término mediante una descripción sino –cuando la referencia es exitosa– con un acto de bautismo fruto de cierta interacción entre el usuario del nuevo término y la entidad que se instituye como su referente. De ese modo, el DCT puede ser caracterizado como una modificación legítima de la teoría causal.<sup>4</sup>

La segunda distinción importante divide aguas entre quienes ven en el DCT una teoría híbrida, por lo que resulta mucho más central respecto del objeto de estas páginas. El punto de quiebre está en el tipo de propiedades que deben ser incluidas en  $D(x)$ . Mientras que para algunos (Nola 1980; Psillos 2012b; Sankey 1997)  $D(x)$  ha de limitarse a consignar

.....  
 4 Aunque su planteo es algo diferente, Psillos parece coincidir con este diagnóstico (*cf.* 2012b 223).

las propiedades relevantes al rol causal que la teoría adjudica a  $x$ , otros (Enc 1976; Psillos 1999) señalan que  $D(x)$  debe incluir una descripción de las propiedades constitutivas de la clase a la que  $x$  pertenece. Es interesante remarcar que Psillos ha defendido ambas variantes en diferentes momentos de su pensamiento. Dado que la primera de ellas será objeto de la cuarta sección, comenzaré por la última.

En la primera versión de su descriptivismo causal, Psillos retoma algunas intuiciones de Enc respecto del modo en que se introducen nuevos términos teóricos, especialmente en su opinión de que la postulación de una entidad inobservable coincide con el descubrimiento de una nueva clase natural, según Enc:

si el científico cree que se trata de una *nueva* clase de objeto, puede introducir un nuevo término en el vocabulario. Ahora el peso de la referencia para el término lo llevarán las propiedades constitutivas de clase atribuidas al objeto y el mecanismo explicativo desarrollado en la teoría, pero no expresiones como “lo que sea que resulte responsable de los fenómenos”. (1976 271)

Dicho en otros términos, la determinación de  $D(x)$  para una nueva entidad postulada implica detallar sus propiedades constitutivas de clase y describir el mecanismo por el cual cumple el rol causal pretendido.

Entiendo que dos cuestiones se alzan como problemáticas para esta posición: en primer lugar, la de cómo es que se llega a constituir la  $D(x)$  adecuada, toda vez que  $x$  puede jugar roles causales diferentes en diversas circunstancias y ser un miembro legítimo de más de una clase natural. En segundo lugar, cómo es que  $D(x)$ , así constituida, es capaz de brindar un relato satisfactorio acerca de la referencia de los términos teóricos.

Respecto de la primera cuestión, el hecho de que lo que constituye una descripción causal adecuada de una entidad depende completamente del contexto teórico en el que se la está postulando, resulta en que algo que pueda ser una propiedad causalmente relevante en un contexto podría bien no serlo en otro. Podría pensarse entonces que lo que debe ser incluido en la  $D(x)$  de una entidad no son las descripciones de propiedades causalmente relevantes respecto de cierto contexto privilegiado (o simplemente considerado), sino respecto de cualquier contexto posible, es decir, de propiedades causalmente relevantes en general. Pero la distinción entre propiedades causales y no causales (¿categorías?, ¿causalmente inertes?) está lejos de poder establecerse sin inconvenientes. De hecho, hay quienes pensamos que el problema consiste justamente en la dudosa posibilidad de hacer tal distinción. Como fuere, Psillos debe hacer mucho más que solo mencionarla para constituir la como una base sólida para su teoría semántica. Además (y este es el punto más serio),

si la  $D(x)$  de una entidad inobservable postulada debiera dar cuenta de todos aquellos aspectos de la entidad capaces de una interacción causal con cualquier otra –pongamos por caso, con nuestros instrumentos de detección–, incluso dando por sentado que esa descripción no agotara todo lo que de la entidad pueda decirse, es por lo menos dudoso que pueda coincidir con la  $D(x)$  de una entidad postulada por una teoría posterior, de modo que el objetivo realista de la continuidad referencial quedaría puesto en jaque. Por otra parte, la determinación de las propiedades constitutivas de clase no resulta menos problemática. La postulación de nuevas entidades inobservables no tiene por qué coincidir con la de una nueva clase natural única, dichas entidades, por el contrario, pueden pertenecer a más de una clase. De ser así, ¿cuáles son las propiedades que han de incluirse en  $D(x)$ ? Ciertamente la respuesta ha de señalar aquellas propiedades que garanticen la continuidad referencial. Para ello sería ventajoso que las propiedades esenciales de las nuevas entidades coincidieran con las constitutivas de clase. Pero esa coincidencia no es en absoluto obvia y los defensores de esta rama del DCT no ofrecen ningún argumento al respecto.

Pero incluso haciendo a un lado las dificultades señaladas respecto de las propiedades consignadas en  $D(x)$ , y suponiendo que la selección adecuada sea posible, resta la espinosa cuestión de hacer de esa descripción una garantía de la continuidad referencial. El mero hecho de que la distinción entre propiedades causales y no causales pudiera hacerse, del mismo modo que suponer que contamos con un criterio pertinente para incluir las propiedades esenciales o constitutivas de clase en  $D(x)$ , no es suficiente para el éxito del descriptivismo causal. Esto se debe a que a la hora de determinar lo que cierta entidad *hace* podríamos prescindir de descripciones que informen acerca de lo que esa entidad esencialmente *es*, o *a qué clase pertenece*.

Imaginemos un mundo en el que las sillas no son artefactos, sino entidades naturales; por si fuera poco, en este particular universo los clavos y los tornillos son objetos inobservables, es decir, su modo de interactuar con nuestro aparato perceptual es tal que no podemos detectarlos a ojo desnudo. Los científicos de este extraño mundo han inferido con notable agudeza que las partes de esas sillas están unidas por clavos, y los han descrito como entidades sólidas, alargadas, con una punta chata y otra aguda, etc. Están orgullosos de sus avances, y con toda razón, sin embargo, se equivocan, puesto que la naturaleza (o algún demiurgo con dotes de carpintero) ha fabricado las sillas con tornillos. Nuestros fantásticos hombres de ciencia han descrito con exactitud todas las propiedades en virtud de las cuales los clavos mantendrían unidas a las partes de las sillas, que son las mismas gracias a las cuales los tornillos hacen lo propio. Con todo, los clavos no son tornillos,

y cuando dicen “clavos” no están refiriéndose en realidad a tornillos. Dicho de otro modo, si la condición para pertenecer  $D(x)$  es ser una descripción de un aspecto relevante al modo en que la entidad postulada juega su rol causal, y si serlo no requiere dar cuenta de la naturaleza de dicha entidad, entonces,  $D(x)$  podría no contener descripciones de las propiedades esenciales constitutivas de la clase a la que la entidad postulada pertenece. Así, el satisfacer  $D(x)$  no es condición suficiente para ser su referente. Lo que aquí se requiere es un argumento que pueda demostrar que las propiedades causalmente relevantes de una entidad inobservable son necesariamente sus propiedades esenciales. Y ese no parece ser el caso. Por el contrario, parece posible que hubiera tanto (i) propiedades causales que no son esenciales, como (ii) propiedades esenciales que no son causales (o al menos, que no determinan cierto rol causal). Lo primero no tiene que ser fatal para el descriptivismo causal, en tanto prescriba que  $D(x)$  debe incluir, además descripciones de las propiedades de tipo (i), descripciones de las propiedades que sí son esenciales. Sin embargo, (ii) presenta un problema mucho mayor. No solo no es obvio que una descripción que dé cuenta del modo en que una entidad inobservable cumple su rol causal deba incluir a todas sus propiedades esenciales, sino que no se brinda ningún argumento positivo acerca de cómo sería posible asegurarse de que toda vez que identifiquemos al conjunto de las propiedades que determinan el rol casual de una entidad en el contexto en el que se la postuló, entonces, el conjunto de sus propiedades esenciales es un subconjunto (propio o impropio) de aquel.

Psillos no contempló de manera explícita estos puntos y ningún crítico llamó la atención sobre ellos, sin embargo, en un reciente trabajo –Psillos (2012b)– brindó una nueva versión de su posición que coincidía con las intuiciones de otros defensores del DCT.

### **Descriptivismo causal, acto II: propiedades causales**

Un modo alternativo (y popular) de interpretar el DCT requiere abandonar la pretensión de que el aspecto descriptivista de la teoría quede ligado a las propiedades constitutivas de clase.<sup>5</sup> Según Psillos,  $D(x)$  puede incluir “*cualquier* propiedad identificatoria estable que contribuya al rol causal atribuido a la entidad postulada *vis-à-vis* a los fenómenos  $\Phi$ ” (2012b 227). Entonces, las *propiedades estables contingentes* ligadas al perfil causal de la entidad postulada son las que llevan

5 Esta versión ha sido defendida por Nola (1980), Sankey (1997) y Psillos (2012b). Dado que este último desarrolló más cuidadosamente los pormenores de esta posición, esta sección se centrará en el análisis de sus argumentos.

el peso de la referencia. Este relato pretende ser no solo coherente con nuestras creencias acerca de términos como “flogisto”, sino acomodarse además a episodios históricos de naturaleza un poco diferente. El ejemplo escogido por Psillos es “clorina”, un término introducido por Humphry Davy en 1812 para nombrar lo que se creía era ácido muriático desflogistizado, previamente aislado por Carl Wilhelm Scheele en 1774. El punto central de esta historia es que Davy, convencido de que su gas era un elemento en lugar de un compuesto, y de que la teoría del flogisto defendida por Scheele estaba errada (puesto que entendía que el gas carecía en realidad de oxígeno), introduce el término “clorina”. La moraleja que Psillos encuentra es que la introducción del nuevo término no corresponde aquí a la postulación de una nueva clase (Davy daba por cierto que había aislado el mismo gas que Scheele), sino al refinamiento de la referencia por vía de las propiedades causales asociadas, es decir, por las descripciones en  $D(x)$ . No obstante, no es en absoluto claro en qué sentido este relato contribuye a mostrar cómo el descriptivismo causal da cuenta de las continuidades o rupturas referenciales adecuadas en la historia de la ciencia. Después de todo, aun admitiendo que Davy y Secheele se interesaron por el mismo gas (*i. e.* por la sustancia responsable de los mismos efectos) no diríamos que la expresión “ácido muriático desflogistizado” refiere a clorina, sino más bien que no creemos hoy que el ácido muriático desflogistizado existe. En cualquier caso, y más allá de si Psillos consigue o no su objetivo con el ejemplo, el punto es que en esta nueva versión de su teoría las propiedades estables contingentes mediante las que se fija la referencia constituyen también propiedades de *identificación*, que permiten seguir el rastro de la referencia a lo largo de los sucesivos ajustes teóricos.

Pero el éxito de esta versión del DCT depende de identificar sin más las propiedades estables contingentes (que podrían llegar a ser constitutivas de clase como no serlo) con las propiedades causales que permiten la detección del referente. Pero, ¿por qué ha de establecerse tal identidad? ¿Por qué las propiedades estables de una entidad deben coincidir con las que determinan su rol causal respecto de los fenómenos que ameritaron su postulación y, además, con las que permiten su identificación en contextos diferentes? Si las propiedades en virtud de las cuales una entidad inobservable juega un rol causal respecto de ciertos fenómenos son consideradas contingentes –puesto que sin duda pueden serlo–, entonces, se vuelven insuficientes como herramientas para fijar unívocamente la referencia. Hay de hecho un problema en el modo mismo de caracterizar la extensión de  $D(x)$ . La mera expresión “propiedades estables contingentes” refleja las exigencias que Psillos pretende alcanzar

sin éxito: *estables* intenta tener la fuerza de *esenciales* o *constitutivas de clase*,<sup>6</sup> el elemento de cohesión que garantice la continuidad referencial, mientras que *contingentes* trae a la expresión el componente falibilista, mostrando que algunas descripciones en  $D(x)$  pueden cambiar sin que la referencia a una y la misma clase natural se altere (Psillos 2012b 227). Pero ese equilibrio es en realidad ficticio: la pretensión de estabilidad no solo es infundada, sino lo suficientemente débil como para no garantizar la continuidad aun en casos en los que los criterios prescritos por el descriptivismo causal se cumplen. El problema general para el DCT termina siendo que las *constitutivas de clase* resultan demasiado y las *propiedades estables contingentes* demasiado poco.

Psillos parece considerar algún aspecto de este problema anticipando dos posibles objeciones a su posición. En primer lugar, puede pensarse que en contextos en los que el referente no posea las propiedades contingentes de detección que se le adscriben, la referencia puede todavía ser exitosa dada la conexión causal entre el término usado y la entidad que denota. De modo semejante, aun suponiendo que el referente posee de manera estable ciertas propiedades contingentes, nosotros podríamos haber equivocado nuestra descripción de esas propiedades. Si eso ocurre, pero estamos todavía en contacto causal con el referente, la referencia podría, sin embargo, ser exitosa. Ambos casos tienden a mostrar que el componente causal es el que realmente carga con el peso de la referencia. La respuesta de Psillos a estos problemas consiste en destacar que ambos elementos, el causal y el descriptivo, intervienen por igual en la fijación de la referencia: “el contacto causal ancla las descripciones al referente: las descripciones permiten rastrear el referente” (Psillos 2012b 227).

Este punto es el principal de los pretendidos beneficios de adoptar esta posición, y su discusión es el objetivo central de esta sección. Psillos considera, de hecho, que son cinco las ventajas más notables del DCT (cf. 2012b 229):

- (1) No hace inevitable la continuidad referencial.
- (2) Explica tanto el fallo como el éxito referencial de un modo no trivial.
- (3) Permite una determinación adicional del referente mediante el agregado de nuevas descripciones.
- (4) Deja abierta la posibilidad de re-identificar la referencia de un término, por ejemplo, cuando resulta que  $x = y$  a pesar de que  $D(x)$  es diferente de  $D'(y)$ .
- (5) Muestra cómo  $D(x)$  es capaz de rastrear la referencia del término en cuestión.

6 En razón del mencionado problema de la constitución de  $D(x)$  a partir de propiedades esenciales, me referiré de aquí en adelante solamente a *propiedades constitutivas de clase*.



Aunque, como acabo de señalar, mi objetivo central es atacar (5), argumentaré contra estos puntos en orden. La primera supuesta ventaja es en realidad extraña. La inevitabilidad de la continuidad referencial es una consecuencia indeseable de la teoría causal pura, por lo que (1) es algo esperable de una teoría que para muchos es considerada eminentemente descriptivista. Por su parte, (2) es en realidad una conjunción de dos proposiciones, y la primera de ellas afirma exactamente lo mismo que (1). La segunda proposición –el verdadero contenido de este punto– afirma que el fracaso referencial no es explicado de modo trivial. Asumo que lo que Psillos tiene en mente al mencionar la *trivialidad* es que no todo cambio en las descripciones asociadas a un término acaban en un cambio de referencia. Nuevamente, eso es natural si el peso de la referencia cae solo sobre un conjunto privilegiado de descripciones, puesto que solo un cambio en esas descripciones acaba en un cambio de referencia. Es por ello que (1) y (2) resultan ventajas en un sentido muy laxo. Ambas se siguen trivialmente del diseño de la teoría, que precisamente fue construida para evitar esas dificultades. En cualquier caso, su carácter de ventajas está supeditado al hecho de que la teoría brinde un relato conceptual e históricamente adecuado de la referencia de los términos teóricos, *i. e.* a la plausibilidad de los puntos siguientes. Prácticamente lo mismo ocurre con (3): dejar la referencia en manos de un tipo especial de descripciones permite que, si efectivamente la referencia se fija a partir de esas descripciones, sucesivas adiciones de descripciones complementarias permitan acrecentar el conocimiento disponible *refinando* la referencia. Claro que, nuevamente, el hecho de que esto constituya una ventaja depende de que la referencia sea efectivamente fijada por las descripciones mencionadas en primer lugar.

El punto (4) representa una ventaja en sentido estricto. La re-identificación del referente es posible en esta versión del DCT en tanto  $D(x)$  no precisa contener descripciones de las propiedades constitutivas de clase de  $x$ . De ese modo,  $D$  y  $D'$  pueden ser sustancialmente diferentes y aun así señalar a las mismas entidades según se comporten causalmente respecto de fenómenos  $\Phi$  y  $\Phi'$  (es importante notar que debe tratarse de *diferentes* fenómenos, dado que asumiendo que  $x$  causa  $\Phi$  de modo regular, las descripciones de las propiedades causales correspondientes no podrían diferir si se tratase de fenómenos de la misma naturaleza). Pero este relato presenta el problema de que el contenido respectivo de  $D(x)$  y  $D'(y)$  en relación con fenómenos  $\Phi$  y  $\Phi'$  *nunca* es suficiente para establecer que  $x = y$ , a menos que incluya descripciones de las propiedades constitutivas de clase correspondientes a  $x$  y  $y$ . Pero dado que esas descripciones son prescindibles, la mencionada re-identificación no resulta posible en todos los casos, y (4) no es, por tanto, la ventaja que aparentaba ser.

Arribamos entonces a (5): si bien el componente causal es importante, las descripciones asociadas a un término son las que permiten rastrear la referencia. Sin embargo, ese no es realmente el caso. Supongamos que en un momento  $t_1$  ante la presencia de los fenómenos  $\Phi$  se postula una entidad  $e$  como responsable de  $\Phi$ , en virtud de ciertas propiedades causales contingentes recogidas en una descripción  $D$ , de modo tal que “ $e$ ” pasa a denotar la entidad  $e$ . Esa entidad, no obstante, posee ciertas propiedades constitutivas de clase  $P$  que no están consignadas en  $D$  –lo que es posible según esta versión debilitada del descriptivismo causal–. En un momento  $t_2$ , ante una nueva instancia de  $\Phi$ , vuelve a usarse el término “ $e$ ” para referir a la misma entidad  $e$  cuya referencia fue fijada por medio de la descripción  $D$ . Sin embargo, en este segundo momento nos encontramos ante una entidad  $f$  que causa  $\Phi$  y satisface  $D$  pero que posee propiedades constitutivas de clase  $Q$  no incluidas en  $D$ . En pro de la argumentación demos por sentado que ambas entidades satisfacen igualmente bien  $D$ , y que es precisamente por aquellas propiedades descritas en  $D$  que causan  $\Phi$ . Sin embargo, “ $e$ ” en  $t_2$  no refiere a  $e$ , sino a  $f$ . Por lo tanto, no es cierto que  $D(x)$  permite rastrear el referente. Si las propiedades de identificación y las propiedades constitutivas de clase no coinciden puede suceder que dos clases podrían compartir las propiedades de identificación y causar los mismos fenómenos y aun así ser distintas en virtud de alguna propiedad constitutiva de clase no detectada (o incluso no detectable) no responsable del rol causal de dichas clases. Tal vez pueda replicarse que la existencia de tal propiedad es una posibilidad meramente lógica, lejana a lo que resulta natural o factualmente posible. Sin embargo, dos cuestiones deben considerarse. En primer lugar, que el caso de la nefrita y la jadeíta parece sugerir que es (naturalmente) posible que propiedades manifiestas idénticas correspondan a sustancias esencialmente diferentes, lo que no dista mucho del escenario analizado;<sup>7</sup> en segundo término, debe recordarse que el mismo Psillos –en el contexto de la discusión en torno a la identificación de los objetos en ontologías estructuralistas (cf. 2012a 174)– no reconoce como problemático aceptar el *quidditismo* como modo de diferenciar entidades (¡y las *quidditas* no son siquiera propiedades!).

Psillos tuvo dos grandes aciertos, uno en cada una de sus defensas del DCT, que coinciden con las dos versiones disponibles de esa posición. Esos aciertos, sin embargo, no son suficientes para lograr una teoría de la referencia satisfactoria para las intuiciones realistas. En su primera versión identificó correctamente el sentido en que una teoría híbrida puede incorporar el elemento causal, dando cuenta de las

7 Para un tratamiento riguroso de las cuestiones filosóficas en torno al caso del jade, véase LaPorte (2009).

propiedades constitutivas de clase de la entidad postulada que tienen relación con los fenómenos. Si esas descripciones resultan verdaderas la referencia del nuevo término queda garantizada incluso a la luz de sucesivos cambios y/o adiciones en el resto de las descripciones asociadas. Psillos lo vio claramente entonces: “la apelación a propiedades constitutivas de clase es esencial porque son estas propiedades las que, en última instancia, fijan la referencia del término” (1999 284). Sin embargo, no es posible asegurar que las propiedades responsables del rol causal que una entidad cumple coincidan necesariamente con sus propiedades esenciales, y es por ello que esta teoría fracasa. La segunda versión es capaz de subsanar esa laguna, no apelando ya a propiedades constitutivas de clase sino a propiedades estables contingentes que determinan el rol causal de la entidad postulada. Pero este acierto de su nueva versión reduce la continuidad referencial a una mera expresión de deseo, puesto que la verdad de las descripciones asociadas no resulta suficiente para fijar ni rastrear la referencia. El sentido último del DCT termina entonces por desmoronarse.

### Referencia y oraciones de Ramsey

La enorme variedad de dificultades que presenta brindar una teoría de la referencia para los términos teóricos ha motivado la adopción de estrategias alternativas tendientes a brindar una solución “externa” al problema, o incluso a disolverlo. Sin duda la más popular es la que apela a las oraciones de Ramsey como un modo de eliminar la aparición de términos teóricos mediante una reducción del enunciado completo de una teoría. Psillos se esfuerza por mostrar que cierta versión de este enfoque coincide con los presupuestos del DCT. Esta sección está dedicada a mostrar que ese no es el caso. Persiguiendo ese objetivo crítico, sin embargo, esta última parte del presente trabajo visita problemas que guardan cierta proximidad con el de la referencia de los términos teóricos, por lo que las conclusiones alcanzadas tienen una dimensión de interés adicional.

Antes de abordar dichas cuestiones resulta menester comenzar por una breve exposición de la técnica de análisis de las teorías introducida por Ramsey (1931), como un método para la eliminación de los términos teóricos de la formulación de una teoría. La idea central es la siguiente: asumamos que tenemos una teoría finitamente axiomatizable:

$$(1) \quad T [O_1, O_2, \dots, H_1, H_2, \dots]$$

Donde  $O_1, O_2, \dots$  son términos observacionales primitivos (constantes de individuo, constantes de predicado, etc.) y  $H_1, H_2, \dots$  términos teóricos primitivos. Dado que es finitamente axiomatizable podemos pensarla como una única proposición compleja. La propuesta de Ramsey

consiste en generalizar existencialmente sobre (1) con el propósito de tener como resultado una oración de la forma

$$(2) (\exists x_1) (\exists x_2) \dots T [o_1, o_2, \dots, x_1, x_2, \dots]$$

En (2) todos los términos teóricos son reemplazados por variables  $x_1, x_2, \dots$  ligadas por cuantificadores existenciales de segundo orden,<sup>8</sup> por lo que al menos nominalmente dichos términos son eliminados. El resultado es lo que se conoce como “oración de Ramsey”. Como el mismo Ramsey destacó, todas las consecuencias lógicas de (1) que contienen solo términos  $o_1, o_2, \dots$  se siguen también de (2), por lo que, si estamos interesados únicamente en las consecuencias observacionales de  $T$ , podemos deshacernos del vocabulario teórico sin pérdida, reemplazándola por su oración de Ramsey. Además, como Carnap (1936) hizo notar, si  $T$  es la teoría presentada en (1) y  $T^R$  su oración de Ramsey construida al modo de (2),  $T$  es equivalente a la conjunción  $T^R \& (T^R \rightarrow T)$ . El segundo término de dicha conjunción es conocido como oración de Carnap, y afirma que *si* existe una clase de individuos que satisface la oración de Ramsey, *entonces* los términos teóricos de  $T$  (*i. e.*  $H_1, H_2, \dots$ ) denotan esa clase. De ese modo, para Carnap,  $T^R$  captura el contenido empírico de la teoría, mientras que  $T^R \rightarrow T$  hace lo propio con su contenido analítico (y todo ello es equivalente a  $T$ ).

Hay más de un modo en que esto se ha interpretado en cuanto concierne al significado de los términos teóricos. Psillos distingue adecuadamente dos tendencias generales al respecto que llama *eliminativista* y *vindicatoria* (*cf.* 2012b 230). La primera considera que dicha equivalencia muestra el carácter prescindible de los términos teóricos, por lo que el problema de su referencia se desvanece por completo. La vindicatoria juzga, en cambio, que el problema queda solucionado, en cuanto lo que se muestra es que la teoría define implícitamente sus términos teóricos. Psillos rechaza de plano la primera de esas interpretaciones, y está en lo cierto al hacerlo. Las razones son complejas y se vinculan con un viejo problema que se ha manifestado de más de una manera en filosofía. En términos sencillos la cuestión es la siguiente. Asumamos con la interpretación eliminativista que  $T^R$  reemplaza a  $T$  sin pérdida. Agregando tan solo suposiciones básicas sobre la cardinalidad de sus dominios resulta que si  $T$  es empíricamente adecuada, entonces, es verdadera. Intuitivamente: dado que  $T^R$  es consistente, tiene un modelo, llamémoslo  $M$ . Sea  $w$  el modelo pretendido de  $T$  y supóngase que  $M$  y  $w$  tienen la misma cardinalidad. Dado que  $T$  es empíricamente adecuada, tanto ella como  $T^R$  salvarán los mismos fenómenos, *i. e.* el submodelo

8 Como se verá, vale notar que (2) es una oración de segundo orden y no una de primer orden como lo es (1).

empírico de  $T^R$  será idéntico al de  $T$ . Ahora bien, dado que  $M$  y  $w$  tienen la misma cardinalidad, puede definirse una relación biunívoca  $R$  entre los dominios de  $M$  y  $w$  preservando toda la estructura definida por las relaciones teóricas en  $w$ . Puesto que  $R$  garantiza la isomorfía entre ambos dominios,  $T^R$  es verdadera en cuanto tiene un modelo en el dominio pretendido de la teoría; y si es cierto que  $T^R$  reemplaza a  $T$  sin pérdidas, entonces,  $T$  es verdadera.<sup>9</sup> En pocas palabras, si  $T^R$  es empíricamente adecuada, siempre pueden definirse trivialmente relaciones en el mundo que la satisfagan por completo, por lo que si realmente expresa todo lo relevante en  $T$ , la interpretación eliminativista parece colapsar en algo muy semejante al empirismo constructivo de van Fraassen. El punto central de esta argumentación aparece por primera vez en la crítica que Newman (1928) hace al empirismo de Russell, y que se ha convertido en una objeción clásica contra muchas formas de realismo estructural –es decir, contra la tesis de que la ciencia solo conoce la estructura del mundo inobservable–.<sup>10</sup> Asimismo, como van Fraassen ha señalado, la radicalización de ese mismo punto es lo que se encuentra en la base del llamado “argumento modelo-teórico” que Putnam (1977) presentó contra el realismo metafísico (cf. van Fraassen 2008 230). Pero más allá de suscribir las razones de Psillos contra el enfoque eliminativista, creo que hay motivos aún más profundos para rechazarlo. El eliminativismo, después de todo, no es tan eliminativista como aparenta ser. Claro que la oración de Ramsey prescinde de los términos teóricos, de hecho, eso es exactamente lo que hace. Pero la mera afirmación existencial de las propiedades que esos términos designaban determina restricciones a las clases de modelos que pueden satisfacer dicha oración. Como Hintikka (1998) mostró convincentemente, los *conceptos teóricos* sobreviven –en un sentido mínimo pero cognitivamente relevante– dentro de la oración de Ramsey en el significado de sus cuantificadores de segundo orden.

Hay un modo obvio en el que el enfoque vindicatorio elude algunas consecuencias del problema de Newman. La equivalencia entre  $T$  y la conjunción  $T^R \wedge (T^R \rightarrow T)$  se interpreta ahora solo como una forma de hacer patente el modo en que  $T$  define implícitamente sus términos teóricos, por lo que la oración de Ramsey ya no es un reemplazo para la teoría original y la adecuación empírica no colapsa en verdad. Pero ese no es el fin de la historia. Incluso con esa modificación, el sentido de esa definición implícita de los términos teóricos de  $T$  depende de la oración de Carnap (*i. e.* del condicional  $T^R \rightarrow T$ ), y lo que esta dice es que las entidades (no observacionales) que satisfagan  $T^R$  constituyen

9 Para una prueba completa véase Ketland (2004).

10 Ainsworth (2009) reconstruye cuidadosamente las discusiones en torno a la objeción de Newman.

los referentes de los términos teóricos de  $\tau$ . Pero el caso es que podría haber más de una clase de entidades que cumplan dicha condición; en otras palabras,  $\tau^r$  podría realizarse en múltiples clases. Como Psillos correctamente reconoce, el auténtico problema es la *realizabilidad múltiple* (cf. 2012b 232). Lewis había identificado mucho antes el problema y había concluido que el único caso en el que puede considerarse que se consigue una definición implícita de los términos teóricos de la teoría de base es cuando su oración de Ramsey tiene una única realización (cf. 1970 432). En un trabajo posterior, en el contexto de la discusión del argumento modelo-teórico, este autor señala que la única manera de garantizar una única realización para las teorías es poner restricciones causales a la referencia de sus términos (cf. Lewis 1984). Su elección para estas restricciones se inclina por la propuesta de Merrill (1980) acerca de lo que considera debe ser una respuesta realista al problema presentado por Putnam (sin ser él mismo un realista). En opinión de Merrill, el foco del impacto del argumento modelo-teórico se encuentra en una caracterización incompleta del realismo metafísico, puesto que este no debe incluir solo la creencia de que el mundo es nada más que una clase de entidades, sino además de que “*esas entidades mantienen unas con otras ciertas relaciones objetivas*” (Merrill 1980 72). La única interpretación verdadera de una teoría es aquella que captura las relaciones *reales o naturales*, las relaciones que hacen del mundo un *dominio estructurado* de entidades. Psillos hace suya esta estrategia, señalando que el problema de la realizabilidad múltiple muestra que “es necesario abandonar los relatos puramente descriptivistas acerca de la fijación de la referencia” (2012b 233). Al igual que Lewis, considera que el camino a seguir es el descriptivismo causal:

Desde un punto de vista realista, una restricción importante para la teoría de la referencia [...] es que pueda acomodar las prácticas lingüísticas a la estructura causal del mundo [...], el enfoque descriptivista causal de la referencia de los términos teóricos está en condiciones de lograr un equilibrio entre, por un lado, la dependencia de los términos teóricos respecto de la teoría y, por otro lado, el ver a dichos términos como transteóricos de un modo relevante. (Psillos 2012b 233)

Pero el movimiento desde el problema de la realizabilidad múltiple al descriptivismo causal es al menos apresurado. En primer lugar, debe destacarse que el contexto en el que Lewis adopta tal estrategia es suficientemente diverso. El problema de Newman tiene en el marco del argumento modelo-teórico algunas aristas diferentes a las que presenta en su aplicación a la interpretación de las oraciones de Ramsey. Para empezar, el argumento modelo-teórico adiciona un resultado meta-matemático, el teorema Loewenheim-Skolem-Tarski-Vaught,

que muestra que si una teoría tiene un modelo infinito entonces tiene modelos de cada cardinalidad infinita; ello permite inferir que cualquier teoría consistente es (trivialmente) verdadera acerca del mundo. Pero –y esto es lo más importante– el argumento de Putnam asume una premisa ausente en el marco del análisis de Ramsey, a saber, que todos los términos de una teoría están igualmente sujetos a cambios radicales de significado. El enfoque de Ramsey, por el contrario, considera el vocabulario observacional como no problemático operando una reducción (sea eliminativista o vindicatoria) sobre el resto de los términos no lógicos de la teoría. Las restricciones causales de Lewis y Merrill permiten garantizar la unicidad de la realización, pero como el mismo Lewis reconoce, no son el único camino posible, ni siquiera el más transitado. El anclaje de ciertos significados al mundo por vía de una relación causal entre términos y sus denotaciones, aunque más no fuere en algunos casos, es todo lo que se necesita para evitar las conclusiones de Putnam. Y *algunos casos* es la expresión clave de la última afirmación, pues esos casos corresponden a la referencia de nombres y términos de clase que designan entidades observables. Después de todo no hay aquí ninguna novedad, dado que (como acabo de señalar) esa referencia no problemática se da por sentada ya en la oración de Ramsey cuando se mantienen las constantes  $o_1, o_2, \dots$  luego de la reducción. En otras palabras, la teoría causal de la referencia es suficiente para bloquear las implicaciones escépticas del argumento modelo-teórico contra el realismo. La elección de Lewis, en este caso, obedece a motivos independientes: “Dada mi preferencia por el descriptivismo causal –que es de hecho solo más descripciones, solo más teoría– debo buscar en otra parte [fuera de las teorías causales puras] mi restricción salvadora” (1984 227). Esa es la razón por la que adopta el realismo metafísico de Merrill. Pero esta cita es esclarecedora en un sentido adicional, que revela las razones del fracaso del proyecto de Psillos: el descriptivismo causal no es por suficiente para determinar una única interpretación de las teorías  $o$ , en otras palabras, para fijar unívocamente la referencia de los términos teóricos. Y lo que en el fondo implica el hecho de que el descriptivismo causal sea *solo más descripciones, solo más teoría* –como agudamente apunta Lewis– es que, a diferencia del enfoque vindicatorio sobre las oraciones de Ramsey, no puede escapar al problema de la realizabilidad múltiple. Eso es lo que, en otros términos, mostró la sección anterior de este trabajo. Como allí fue señalado, el hecho de que una clase cumpla con los requisitos fijados por el descriptivismo causal para ser el referente de un término –ser la causa de los fenómenos  $\Phi$  y satisfacer la descripción  $D(x)$ –, no resulta condición suficiente para que en efecto sea el referente de ese término. La razón de ello es, en última instancia, que las propiedades casuales contingentes descritas en  $D(x)$

y su relación causal con  $\Phi$  pueden corresponder a clases con diferentes propiedades esenciales, *i. e.* pueden estar *múltiplemente realizadas*. El enfoque vindicatorio acerca de las oraciones de Ramsey no pretende una fijación explícita de la referencia de los términos teóricos, sino por el contrario una definición implícita de dichos términos, por lo que la adición de un presupuesto metafísico sobre el carácter naturalmente estructurado del mundo es suficiente para escapar del problema de la realizabilidad múltiple: un término teórico referirá a *cualesquiera* sean las propiedades esenciales de la única clase que satisface la oración de Ramsey en la que figura.

Esta característica le aporta a la estrategia de dar definiciones implícitas de los términos teóricos por medio de oraciones de Ramsey dos ventajas adicionales. La primera se desprende de lo dicho hasta aquí: partiendo de un compromiso más débil que la semántica realista de Psillos presupuesta en su DCT, la estrategia de Ramsey permite arribar a un resultado más preciso, y por tanto más fuerte, respecto de la fijación de la referencia de los términos teóricos. El hecho de que la oración de Ramsey sea una fórmula de segundo orden implica que se hace abstracción de la naturaleza concreta de las propiedades y relaciones inobservables postuladas por la teoría de base. De ese modo, el defensor de la estrategia vindicatoria, a diferencia del descriptivista causal, no alega un conocimiento positivo respecto de dichas propiedades, siendo capaz pese a ello de dar cuenta de la referencia de los términos teóricos en su marco realista. En otros términos, pese a que sus compromisos son *epistémicamente más débiles*, sus resultados son *semánticamente más fuertes*.

La segunda ventaja se vincula a la compatibilidad de la estrategia de Ramsey con una variante ya mencionada del realismo científico: el realismo estructural. La idea de que todo nuestro conocimiento del mundo inobservable es puramente estructural se ha expresado frecuentemente –por ejemplo en Maxwell (1970) y Worrall y Zahar (2001)– en términos de que dicho conocimiento puede ser expresado por medio de la oración de Ramsey de una teoría aceptada. La idea es que la ramificación de teorías nos brinda un modo de aislar los compromisos puramente estructurales del resto de bagaje teórico. De este modo, el realismo estructural se presenta como una posición capaz de conciliar las intuiciones que motivan los dos argumentos antagónicos que definen la polémica, el *argumento del no-milagro* y la *meta-inducción pesimista*. Sumariamente, la permanencia de los componentes estructurales a lo largo del cambio teórico explica el creciente éxito predictivo de la ciencia, mientras que las variaciones en los compromisos ontológicos dan cuenta de las razones por las que las teorías pasadas fueron abandonadas. De este modo, la estrategia vindicatoria resulta ventajosa en



un sentido doble. Por una parte, la idea de que son las oraciones de Ramsey las que aportan la medida del compromiso epistémico y semántico que debemos asumir se ve respaldada por los desarrollos históricos y conceptuales que durante las últimas tres décadas han respaldado al realismo estructural. Pero, además, dicha posición puede encontrar en la estrategia vindicatoria un modo de completar sus afirmaciones epistémicas con una teoría de la referencia adecuada. En cualquier caso, las relaciones del enfoque aquí defendido con el realismo estructural han de ser precisadas en futuras investigaciones.

El descriptivismo causal pretende, por su parte, que la apelación a propiedades causales contingentes sea suficiente para fijar y rastrear la referencia de un término. Pero eso, como hemos visto, no es el caso, y recurrir a un presupuesto metafísico como el mencionado no colaboraría de modo alguno para que lo fuera. Incluso si el mundo estuviese naturalmente estructurado las propiedades contingentes podrían realizarse sobre múltiples clases, por lo que no resultarían suficientes para fijar la referencia. Tal vez una de las lecciones más importantes de esta historia –lección que, curiosamente, muchos aprendimos de Putnam– es que las propiedades constitutivas de clase resultan ser, después de todo, el puerto último donde ha de anclar la referencia.

### Conclusión

Dos de las formulaciones más recientes y acabadas del DCT –que recogen varias intuiciones presentes en versiones previas– fracasan a la hora de brindar un relato acerca de cómo queda fijada y puede rastrearse la referencia de los términos teóricos. La primera, por no poder justificar la identificación entre las propiedades causalmente relevantes en cierto contexto y las que constituyen una clase natural. La segunda, en razón de que las descripciones de propiedades causales contingentes vinculadas a una nueva entidad postulada no resultan suficientes para fijar la referencia. El problema no es otro que el de la realizabilidad múltiple de esas propiedades. La estrategia vindicatoria asociada a las oraciones de Ramsey, por el contrario, supera el problema de la realizabilidad múltiple adicionando un supuesto realista extra respecto del carácter naturalmente estructurado del mundo. Tal vez la moraleja sea que el remedio para el realismo es más realismo. La idea de que la naturaleza puede ser *cortada por sus coyunturas* es casi tan vieja como la filosofía y suena de maravillas a los oídos de los realistas, o al menos a los míos. Sin embargo, la pretensión de ofrecer definiciones implícitas de los términos teóricos no deja de tener problemas intrínsecos, entre ellos que, en última instancia, nunca sabemos exactamente a *qué* nos referimos, por lo que el modo en que la referencia de un término queda fijada sigue siendo casi un misterio. Tal vez la referencia sea el problema,

tal vez solo nuestras ideas al respecto. En cualquier caso, el DCT no es el camino para resolver el problema de los términos teóricos. Esa es una gran noticia: debemos seguir buscando.

## Bibliografía

- Ainsworth, P. M. "Newman's Objection." *The British Journal for the Philosophy of Science* 60.1 (2009): 135-171. <http://dx.doi.org/10.1093/bjps/axn051>
- Bird, A. *Thomas Kuhn*. London: Routledge, 2000.
- Carnap, R. "Testability and Meaning." *Philosophy of Science* 3.4 (1936): 419-471. <http://dx.doi.org/10.1086/286432>
- Enç, B. "Reference of Theoretical Terms." *Noûs* 10.3 (1976): 261-282. <http://dx.doi.org/10.2307/2214608>
- Halvorson, H. "What Scientific Theories Could Not Be." *Philosophy of Science* 79.2 (2012): 183-206. <http://dx.doi.org/10.1086/664745>
- Hintikka, J. "Ramsey Sentences and the Meaning of Quantifiers." *Philosophy of Science* 65 (1998): 289-305.
- Jackson, F. "Reference and Description Revisited." *Noûs* 32.S12 (1998): 201-218. <http://dx.doi.org/10.1111/0029-4624.32.s12.9>
- Ketland, J. "Empirical Adequacy and Ramsification." *The British Journal for the Philosophy of Science* 55.2 (2004): 287-300. <http://dx.doi.org/10.1093/bjps/55.2.287>
- Kripke, S. A. *Naming and Necessity*. Cambridge: Harvard University Press, 1980.
- Kroon, F. W. "Causal Descriptivism." *Australasian Journal of Philosophy* 65.1 (1987): 1-17. <http://dx.doi.org/10.1080/00048408712342731>
- Kroon, F. W. "Theoretical Terms and the Causal View of Reference." *Australasian Journal of Philosophy* 63.2 (1985): 143-166. <http://dx.doi.org/10.1080/00048408512341771>
- Kroon, F. W. "Names, Plans, and Descriptions." *Conceptual Analysis and Philosophical Naturalism*. Eds. David Braddon-Mitchell and Robert Nola. Cambridge: MIT Press, 2008. 139-158. <http://dx.doi.org/10.7551/mitpress/9780262012560.003.0006>
- Kuhn, T. S. *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press, 1962.
- LaPorte, J. *Natural Kinds and Conceptual Change*. Cambridge: Cambridge University Press, 2009.
- Lewis, D. "How to Define Theoretical Terms." *The Journal of Philosophy* 67.13 (1970): 427-446. <http://dx.doi.org/10.2307/2023861>
- Lewis, D. "Putnam's Paradox." *Australasian Journal of Philosophy* 62.3 (1984): 221-236. <http://dx.doi.org/10.1080/00048408412340013>
- Loar, B. "The Semantics of Singular Terms." *Philosophical Studies* 30.6 (1976): 353-377. <http://dx.doi.org/10.1007/BF00372537>
- Lutz, S. "What's Right with a Syntactic Approach to Theories and Models?" *Erkenntnis* 79.8 (2014): 1475-1492.

- Maxwell, G. "Structural Realism and the Meaning of Theoretical Terms." *Analyses of Theories and Methods of Physics and Psychology*. Eds. Michael Radner and Stephen Winokur. Minneapolis: University of Minnesota Press, 1970. 181-192.
- Merrill, G. H. "The Model-theoretic Argument Against Realism." *Philosophy of Science* 47.1 (1980): 69-81.
- Newman, M. H. A. "Mr. Russell's Causal Theory of Perception." *Mind* 37.146 (1928): 137-148.
- Niiniluoto, I. *Critical Scientific Realism*. Oxford: Oxford University Press, 1999.
- Nola, R. "Fixing the Reference of Theoretical Terms." *Philosophy of Science* 47.4 (1980): 505-531. <http://dx.doi.org/10.1086/288954>
- Psillos, S. *Scientific Realism: How Science Tracks Truth*. London; New York: Routledge, 1999.
- Psillos, S. "Adding Modality to Ontic Structuralism: An Exploration and Critique." *Structural Realism*. Eds. Elaine Landry and Dean Rickles. Dordrecht; New York: Springer, 2012a. 169-185.
- Psillos, S. "Causal Descriptivism and the Reference of Theoretical Terms." *Perception, Realism, and the Problem of Reference*. Eds. Athanassios Raftopoulos and Peter Machamer. Cambridge; New York: Cambridge University Press, 2012b. 212-238. <http://dx.doi.org/10.1017/CBO9780511979279.010>
- Putnam, H. "The Meaning of 'Meaning'." *Language, Mind and Knowledge*. Ed. Keith Gunderson. Minneapolis: University of Minnesota Press, 1975. 131-193.
- Putnam, H. "Realism and Reason." *Proceedings and Addresses of the American Philosophical Association* 50.6 (1977): 483-498. <http://dx.doi.org/10.2307/3129784>
- Putnam, H. "Reference and Truth." *Philosophical Papers: Realism and Reason*. Vol. 3. Cambridge: Cambridge University Press, 1983. 69-86. <http://dx.doi.org/10.1017/CBO9780511625275.006>
- Raatikainen, P. "Theories of Reference and the Philosophy of Science." *1st Conference of the European Philosophy of Science Association*. EPSA07, Madrid. 15 November 2007.
- Ramsey, F. P. "Truth and Probability [1926]." *The Foundations of Mathematics and Other Logical Essays*. By: Frank Plumpton Ramsey. Ed. Richard Bevan Braithwaite. London: K. Paul, Trench, Trubner and Co., 1931. 156-198.
- Sankey, H. "Incommensurability." *Theoria. Revista de Teoría, Historia y Fundamentos de la Ciencia* 12.3 (1997): 425-445.
- van Fraassen, B. *Scientific Representation: Paradoxes of Perspective*. New York: Oxford University Press, 2008.
- Worrall, J., and Zahar, E. "Ramsification and Structural Realism." *Poincaré's Philosophy: From Conventionalism to Phenomenology*. By: Elie Zahar. Chicago: Open Court, 2001. 236-251.