

## CONCEPTOS TEORICOS Y TEORIAS CIENTIFICAS

*C. Ulises Moulines*

### 0. INTRODUCCION

Las disciplinas científicas se caracterizan, entre otras cosas, por el uso de un vocabulario específico, de ciertas palabras y expresiones que no son del acervo común de los lenguajes comunmente hablados, sino que son introducidas especialmente en un contexto científico. El sentido de tales términos no puede ser apresado plenamente si no se tiene un conocimiento mínimo de la disciplina en la que aparecen. No nos referimos aquí a expresiones procedentes del lenguaje matemático puro (expresiones aritméticas, geométricas o algebraicas, por ejemplo), sino a términos que tienen, o pretenden tener, una referencia en la realidad empírica, pero cuyo manejo adecuado es muy difícil, cuando no imposible, para personas que no estén suficientemente entrenadas en la disciplina en la que aparecen. Ejemplos de tales términos o expresiones, característicos de distintas disciplinas científicas, son: «fotón», «spin», «campo electromagnético», «entropía», «momento angular», «ion», «placa tectónica», «gen», «reflejo condicionado», «plusvalía», «juego de suma cero». Algunos de ellos han hecho ya su entrada en el lenguaje común no-científico, como es el caso de «entropía», «reflejo condicionado» o «plusvalía», pero, incluso en esos casos, su uso por parte de los hablantes no especializados suele ser metafórico, inseguro; en definitiva, el hablante normal es consciente de no ser capaz de usarlos con la misma soltura y propiedad con las que usa los términos usuales de su vida cotidiana, como «agua», «árbol», «montaña», «casa», etc.

La especificidad del primer grupo de términos no consiste meramente en que fueron introducidos en algún momento del desarrollo científico. Muchos términos que deben su introducción al desarrollo científico y, sobre todo, tecnológico, y que por tanto no existían en las lenguas naturales antes de la Revolución científica, son ahora de uso común; tales

son: «gas», «microscopio», «célula», «inflación». Probablemente, muchos de estos términos de origen científico que ahora son del acervo común fueron en su origen igual de difíciles de manejar que los primeros mencionados; y probablemente algunos de éstos acaben por ser tan cotidianos como «gas» o «microscopio». Sin embargo, con independencia del desarrollo histórico de la cultura humana, lo que podemos constatar es que, en cada corte sincrónico en la evolución de la ciencia, nos encontramos con una serie de expresiones que sólo son realmente inteligibles y empleadas de manera pertinente por los expertos en una disciplina dada.

Lo esencial en los términos a los que nos referimos no es pues que tengan un origen científico, sino que su uso sólo puede estar sancionado por una teoría científica, y que sólo quien conozca bien esa teoría, podrá hacer un uso genuino de ellos. Así, quien no tenga idea de termodinámica, no podrá emplear apropiadamente la palabra «entropía»; sólo un geólogo sabe realmente de qué está hablando cuando usa «placa tectónica», y sólo en el contexto de la teoría de juegos tiene un sentido preciso la expresión «juego de suma cero». En consecuencia, es adecuado caracterizar estas expresiones específicas del lenguaje científico como «*términos teóricos*». (En lo sucesivo, se hablará de «*términos teóricos*» cuando se quiera hacer hincapié en ellos en tanto que entidades lingüísticas que aparecen en la formulación canónica de una teoría, y de «*conceptos teóricos*», de un modo más general, cuando la discusión no presuponga una formulación específica.) A todas las demás expresiones referidas a la realidad empírica, que no son términos teóricos, las llamaremos de momento simplemente «*términos no-teóricos*».

La presencia de términos teóricos en las diversas disciplinas científicas es algo más que una mera curiosidad filológica. Ella les plantea al filósofo que reflexiona sobre la ciencia y al científico que se interesa por los fundamentos de su propio quehacer una serie de cuestiones que trascienden el horizonte estrictamente lingüístico. Estas cuestiones son de orden:

—semántico-filosófico: ¿cuál es el significado de los términos teóricos y cómo se relaciona con el significado de los términos no-teóricos?

—epistemológico: ¿son esenciales los conceptos teóricos al conocimiento genuinamente científico? ¿Son responsables de un tipo de conocimiento distinto del ordinario?

—ontológico: ¿existen los referentes de los términos teóricos?; ¿existen en el mismo sentido en que existen los referentes de palabras tales como «agua», «árbol», «casa», etc.?

—metodológico; ¿juegan un papel especial en la capacidad explicativa y predictiva de las teorías científicas?

—metateórico: ¿qué implicación tiene la presencia de términos teóricos para la estructura de las teorías científicas? ¿Son las teorías científicas justamente *teorías* (y no meras compilaciones de datos o de regularidades) por el hecho de estar construidas *con* conceptos teóricos?

Todas estas cuestiones no son en absoluto triviales. Antes bien, han jugado un gran papel en el desarrollo de la filosofía de la ciencia del siglo XX; incluso puede decirse que fueron determinantes en la constitución de la filosofía de la ciencia como disciplina relativamente autónoma a partir de los años treinta. Muchos de los más connotados filósofos de la ciencia de nuestra época han tratado de dar respuestas sistemáticas y detalladas a esas preguntas que, genericamente y con cierto abuso del lenguaje, podemos englobar bajo el rótulo «problema de los términos (o conceptos) teóricos». Recordemos sólo a algunos de ellos (por orden cronológico): Frank P. Ramsey, Rudolf Carnap, Carl G. Hempel, Herbert Feigl, Richard Braithwaite, Ernest Nagel, Hilary Putnam, Wolfgang Stegmüller y Joseph D. Sneed.

No todos los epistemólogos influyentes de este siglo se han ocupado del problema de los términos teóricos. Algunos simplemente lo han ignorado (como Patrick Suppes), otros lo han considerado de importancia secundaria. Así Karl R. Popper y sus discípulos consideran que la distinción entre conceptos teóricos y no-teóricos es sólo una «cuestión de grado», sin especial importancia metodológica o epistemológica; parecida es la opinión de Clark Glymour, aunque difiera de Popper en tantos otros aspectos; finalmente, algunos como Feyerabend consideran que la discusión en torno a los términos teóricos es sólo una entelequia de filósofos, un problema artificial construido por filósofos sin verdadero contacto con la realidad del desarrollo científico.

La opinión de que el problema de los términos teóricos carece de importancia epistemológica, y aún más la idea de que es un artificio de filósofos, es no sólo errónea por razones sistemáticas, sino que revela un desconocimiento pasmoso de la historia de las reflexiones metodológicas emprendidas *por los propios científicos* desde hace tres siglos, como indicaremos a continuación. Por lo demás, el problema de los términos teóricos no es sólo significativo por sí mismo, es decir, por las dificultades que plantea la existencia de tales términos en la ciencia, sino también porque él va íntimamente ligado a otro problema metateórico, quizás el más crucial en la filosofía de la ciencia: la naturaleza de esas entidades que llamamos «teorías científicas». La solución que se dé al problema de los términos teóricos condiciona y es condicionada por la concepción que se tenga de las teorías científicas en general.

## I. SURGIMIENTO DEL PROBLEMA DE LOS CONCEPTOS TEORICOS

La Revolución científica del siglo XVII se puso en marcha bajo el signo de un rechazo radical de la metafísica escolástica. Este rechazo es explícito en los grandes iniciadores de esta revolución (Kepler, Galileo, Francis Bacon, Descartes) y sus discípulos, independientemente de las diferentes filosofías por ellos adoptadas. Se achacaba a la metafísica escolástica (el parangón de la metafísica en general) el haber impedido el genuino co-

nocimiento científico por el abuso de términos abstractos, vacíos de contenido empírico. Incluso un racionalista matemático como Descartes intentó, en su obra sobre mecánica, que sus términos fundamentales (velocidad, tamaño, choque) estuvieran lo más estrechamente asociados posible a lo empíricamente constatable. El rechazo de la metafísica escolástica por parte de los científicos del XVII forma parte de una atmósfera intelectual más general, enemiga de la palabrería hueca y precursora del espíritu iconoclasta de la Ilustración. Recuértese la mofa que hace Molière de la «virtud dormitiva» como pretendido principio para explicar el sueño —el paradigma de hueca explicación escolástica de un fenómeno natural—. Lo que más aborrecía el científico medio del siglo XVII era la introducción de tales «cualidades ocultas», como se las llamaba, en la explicación científica.

Ahora bien, paradójicamente, fue la obra culminante de la Revolución científica, los *Principia* de Newton, la que justamente dio la impresión de volver a introducir las «cualidades ocultas» por la puerta trasera. En efecto, en esa obra jugaba un papel central la noción de fuerza de gravedad (y la noción de fuerza en general), que a los científicos y filósofos más esclarecidos contemporáneos de Newton les había de parecer, por su carácter abstracto, un hermano gemelo de la «virtud dormitiva» de Molière. Los dos contemporáneos quizás más brillantes de Newton —Huygens y Leibniz—, a pesar de reconocer el gran talento matemático del primero, se negaron a admitir el valor físico de su sistema, justamente por las «cualidades ocultas» que parece contener, especialmente con su concepto universal de fuerza. Ambos rechazaron lo que calificaban de «extraña metafísica» (cf. Moulines, 1976).

Aunque la mecánica de Newton acabó por imponerse, no por ello dejaron de provocar inquietud sus fundamentos conceptuales en los físicos de los siglos siguientes. A mediados del siglo XVIII, Jean-le-Rond d'Alembert se propone en su *Traité de Dynamique* proscribir las fuerzas newtonianas, a las que califica de «seres oscuros y metafísicos, que no son aptos más que para difundir las tinieblas en una ciencia que en sí misma debería ser clara» (cf. Moulines, 1975, 35). El propósito no tuvo una realización muy efectiva, pues los físicos siguieron operando, aunque fuera con mala conciencia, con entidades tales como fuerzas de atracción, acciones a distancia, espacio y tiempo absolutos. Más de un siglo más tarde, Ernst Mach, Gustav Kirchhoff y Heinrich Hertz (todos ellos primariamente físicos y sólo secundariamente filósofos) volvieron a la carga para «purificar» los fundamentos de la mecánica de lo que ellos consideraban su lastre metafísico. Mach incluso fue un paso más allá que los demás y propuso eliminar el concepto de masa, reteniendo sólo las nociones cinemáticas relativas a un observador como aquellas dignas de constituir la base conceptual de la mecánica (cf. Moulines, 1975). Lo que estos investigadores reprochaban a conceptos tales como «fuerza», «masa», «espacio absoluto», «acción a distancia» (y también luego «éter», «campo electromagnético» y «átomo») era el no tener una refe-

rencia empírica directamente contrastable, el no referirse a «observables», como hoy diríamos y con ello abrir la puerta a una metafísica que nadie (los científicos en primer lugar) quería tener metida en la ciencia.

Sin embargo, al mismo tiempo era claro que conceptos tales como los mencionados estaban tan enraizados en la formulación de las mejores teorías físicas del momento, que el programa de «eliminación de la metafísica» sólo tenía posibilidades de éxito si se sometía el edificio de la física (y de la ciencia empírica en general) a una transformación radical, a una «reconstrucción racional». Ésta fue la bandera que iban a retomar, ya entrado el siglo XX, algunos de los miembros del Círculo de Viena, sobre todo Carnap. Y así es como nació el problema de los términos teóricos: no como invento caprichoso de filósofos sin relación con la ciencia, sino como programa sistemático alentado por la conciencia de un problema que los físicos habían sentido como propio desde por lo menos las postrimerías del siglo XVII.

## II. EL PROGRAMA REDUCCIONISTA

Para Mach, Hertz y tantos otros de los físicos y filósofos de la física de hace cien años (entre ellos el joven Einstein), un concepto físico era admisible solamente si se refería directamente a alguna entidad observable, o bien era «reducible» a una entidad observable. Cualquier término que no fuera observacional, o reducible a términos observacionales, debía ser eliminado del vocabulario de la física. Este principio metodológico había de tener consecuencias no sólo para la reconstrucción de las teorías ya existentes, sino para la construcción de nuevas teorías —como lo prueba el caso de la teoría especial de la relatividad: en el proceso de definir un concepto puramente observacional (u «operacional», como a veces también se dice) de simultaneidad, se eliminan del discurso científico las nociones newtonianas de espacio y tiempo absolutos.

Pero, ¿qué significa exactamente «reducir» un concepto teórico dado a otros conceptos supuestamente observacionales? La forma a primera vista más plausible de entender semejante reducción es en el sentido de una definición o cadena de definiciones. Si un término «sospechoso de metafísica» es definible, aunque sea a través de cadenas más o menos largas y complicadas, a partir de términos que claramente se refieren a observables, entonces el término en cuestión queda libre de sospecha y puede seguir siendo empleado en el discurso científico; de lo contrario, debe ser desechado.

El programa de «purificación antimetafísica» de las ciencias empíricas dependía, pues, de la noción clave de definibilidad. Esta noción, sin embargo, es más difícil de precisar y aplicar de lo que a primera vista parece. Muchas veces creemos haber propuesto una buena definición que luego, ante un análisis cuidadoso, resulta no ser tal. Ello le ocurrió justamente a Mach. En su artículo de 1868 creyó haber definido el concepto

de masa en función de conceptos puramente cinemáticos; sin embargo, puede mostrarse formalmente que la definición es defectuosa por no ser generalizable en el sentido que pretendía Mach (cf. Suppes, 1957, c. VI). En realidad, sería injusto echarle en cara esta deficiencia a Mach, por cuanto una teoría formal adecuada de la definición no iba a surgir sino ya entrado el siglo XX, sobre todo con los trabajos de Padoa y Leśniewski, que a su vez presuponían la lógica matemática moderna. La realización efectiva del programa de depuración conceptual de los físicos tenía que ir de la mano de la aplicación de las herramientas formales de los lógicos, en especial por lo que concierne a la cuestión de la definibilidad.

Un requisito fundamental para garantizar una definición correcta de un predicado  $P$  en función de otros  $Q_1, \dots, Q_n$  es que se pueda formular un bicondicional generalizado en el que  $P$  aparece solo a la izquierda del bicondicional (el *definiendum*) mientras que a la derecha aparecen  $Q_1, \dots, Q_n$  en cierta combinación lógica. De esta manera garantizamos que todos y sólo los casos de aplicación de  $P$  sean también los casos de aplicación de la combinación de  $Q_1, \dots, Q_n$ . Supongamos, por ejemplo, que consideramos a los predicados relacionales « $x$  es padre de  $y$ », « $x$  es madre de  $y$ » y « $x$  es hermano de  $y$ » como claramente observacionales, mientras que tenemos sospechas sobre la observacionalidad de « $x$  es tío de  $y$ ». Podemos «salvar» la noción de tío definiéndola de esta manera:

Para todos  $x$  e  $y$ :  $x$  es tío de  $y$  si y sólo si existe  $z$  tal que:  $x$  es hermano de  $z$ , y  $z$  es padre de  $y$  o  $z$  es madre de  $y$ .

Abreviado simbólicamente:

$$\forall x \forall y (xTy) \leftrightarrow \exists z (xHz) \wedge (zPy \vee zMy).$$

Con ello podemos decir que hemos *reducido* el predicado T («tío») a los predicados H, P y M («hermano», «padre», «madre»), pues efectivamente la fórmula anterior cumple las condiciones de una buena definición.

El programa de fundamentación conceptual de las ciencias empíricas consistió pues, en su primera fase, en proporcionar tales bicondicionales (por complicados que fueran) para todos los términos problemáticos de las ciencias empíricas —precisamente términos teóricos como «electrón», «entropía», «gen», «plusvalía», etc.—, de modo que cada uno de ellos apareciera como *definiendum* de un bicondicional (o una cadena de bicondicionales), cuyo *definiens* (al final) consistiera de términos indudablemente observacionales. En tal caso, todos los problemas arriba mencionados (semánticos, epistemológicos, ontológicos, etc.) con respecto a los términos teóricos se disolverían al ser reformulables en función de términos observacionales aparentemente no problemáticos —o al menos no más problemáticos que términos cotidianos como «agua», «casa», etc.

Esta disolución de la problemática específica de los términos teóricos, sin embargo, sólo está garantizada si se cumplen dos supuestos: 1) que,

*efectivamente*, para *todos* los términos teóricos presentes en las ciencias podemos formular los bicondicionales requeridos; 2) que hay consenso acerca de cuáles son los «términos indudablemente observacionales» que están en la base de las definiciones. Ninguno de ambos requisitos es evidente. Es más, el desarrollo ulterior de la filosofía de la ciencia iba a mostrar que ninguno de ambos supuestos es realizable o, dicho más cautelosamente (porque no existe ninguna prueba formal al respecto), que no hay ninguna buena razón para pensar que lo son.

Carnap fue el primer autor que se propuso resolver de manera sistemática y efectiva ambos problemas en su ingente obra *La construcción lógica del mundo* (cf. Carnap, 1928). Como lenguaje observacional de base (que en lo sucesivo abreviaremos por «L<sub>o</sub>»), Carnap escogió lo que se denomina un «lenguaje fenomenalista», es decir, un lenguaje cuyo vocabulario no puramente lógico o matemático se refiere a las experiencias de un sujeto. Para ello se inspiró en la idea informalmente esbozada por Mach en su principal obra epistemológica, *Análisis de las sensaciones* (cf. Mach, 1883). El lenguaje fenomenalista de Carnap contiene términos relacionales del tipo «x es una experiencia cromática semejante a y». Con esta exigua base, cuyo carácter observacional parece evidente, Carnap logra definir efectivamente, mediante los instrumentos formales de la lógica, la teoría de conjuntos y la topología, una serie de importantes conceptos psicológicos más teóricos (como «cualidad sensible», «espacio visual», etc.). Sin embargo, la empresa se tambalea en el paso crucial a los conceptos fundamentales de la física. Carnap presentó fórmulas reductivas de términos fundamentales como «punto espaciotemporal», «línea-universo» y otros semejantes en función de su L<sub>o</sub>; pero se equivocó al pensar que esas fórmulas eran auténticas definiciones (cf. Moulines, 1991). Más bien se trata de lo que en la terminología metodológica posterior se denominarían «reglas de correspondencia».

Pocos años después de publicar su *Construcción lógica del mundo*, Carnap abandonó el programa fenomenalista de reducción de los conceptos teóricos e intentó fijar L<sub>o</sub> mediante un vocabulario estrictamente *fisicalista*, es decir, con términos cuyos referentes fueran objetos o procesos macroscópicos ordinarios, ejemplificados paradigmáticamente por lo que encontramos en un laboratorio científico —términos tales como «mesa», «regla», «tubo», «aguja» y operaciones físicas asociadas a ellos—. Este cambio de enfoque no se debió tanto a las dificultades encontradas al tratar de definir términos de la física en función de experiencias perceptuales —dificultades que, a primera vista, podían considerarse meramente técnicas y que podrían resolverse en el futuro—, cuanto porque Otto Neurath convenció a Carnap de que un lenguaje fenomenalista era por principio inadecuado debido a la falta de intersubjetividad de sus referentes. El fenomenalismo parecía llevar a un solipsismo incontrolable.

Dejaremos aquí abierta la cuestión de hasta qué punto la acusación de subjetivismo contra el programa inicial de Carnap estaba realmente justificada. En cualquier caso, Carnap y Neurath esbozaron el programa

fiscalista de reducción en una serie de artículos que publicaron en la revista *Erkenntnis* en los años treinta (cf. también Ayer, 1959). Dentro de un espíritu parecido, aunque con menor rigor lógico, hay que ver el operacionalismo del físico P.W. Bridgman por la misma época. Según éste, el significado auténtico del término teórico «temperatura», por ejemplo, se reduce a la serie de operaciones e indicaciones observables asociadas al aparato llamado «termómetro».

Tanto el fiscalismo de Carnap/Neurath como el operacionalismo de Bridgman fueron muy influyentes en la discusión posterior, y no sólo dentro de la filosofía de la ciencia, sino también como programas para colocar sobre fundamentos sólidos a las disciplinas científicas más «jóvenes». Así, el conductismo de B.F. Skinner en psicología o la lingüística de Bloomfield pueden verse como aplicaciones a disciplinas particulares de esa idea general. Y, a partir de entonces, cuando en la discusión epistemológica se habla de «lenguaje observacional» se suele pensar en el vocabulario sobre objetos macroscópicos de laboratorio en el que pensaban Carnap, Neurath y Bridgman, y no en las «sensaciones» de Mach o las «experiencias» del *Aufbau*.

Ahora bien, tanto el fiscalismo como el operacionalismo han resultado de hecho inviables. El primer golpe rudo lo sufrieron por mano del propio Carnap, en sus investigaciones de *Testability and Meaning* (cf. Carnap, 1936-37). En esta monografía demuestra Carnap que, al menos con los instrumentos formales clásicos, no es posible definir los términos *disposicionales* (es decir, los que se refieren a disposiciones de objetos o sistemas) en función de términos observacionales. Ejemplos típicos de términos disposicionales son: «soluble», «elástico», «conductor», «apareable», «inteligente». Está claro que estos términos no designan entidades o propiedades directamente observables: por ejemplo, si predicamos la solubilidad de un terrón de azúcar, no predicamos una propiedad directamente percible como pueda ser su color blanco o bien su carácter rugoso al tacto. Pero, a primera vista, nada parece más fácil que definir la solubilidad de una sustancia en agua, atendiendo a operaciones y reacciones directamente observables, por ejemplo estableciendo:

«x es soluble (en agua) si y sólo si, en caso de que introduzcamos x en agua, x se disuelve».

Simbólicamente:

$$\forall x(Sx \leftrightarrow (IA x \rightarrow Dx)).$$

Sin embargo, tomemos un pedazo de papel, digamos *p*, que nunca haya sido introducido en agua y que quemaremos antes de que nadie pueda introducirlo en agua. Para este *p* se cumple: no-*IAp*. Por las reglas de la lógica clásica se infiere de ahí que el condicional del *definiens* anterior, «*IAp* → *Dp*», es verdadero, y por tanto también lo resultará el *definiens*



*dum*, o sea, *Sp*. De acuerdo a la definición propuesta, ese pedazo de papel resulta ser un objeto soluble en agua, por el mero hecho de no haber sido introducido nunca en agua. Lo cual, naturalmente, es inadmisibile.

Ésta es una constatación válida para todas las disposiciones. Precisamente porque tales conceptos cubren condiciones hipotéticas que quizás nunca se realizarán o incluso no pueden realizarse por principio, su significado no puede agotarse en una serie de predicados puramente observacionales. Por otro lado, sería absurdo eliminarlos del discurso científico por «metafísicos»: ellos abundan en todas las disciplinas científicas, y muchas teorías bien establecidas no podrían formularse sin ellos.

Ahora bien, no son sólo los términos disposicionales los que causan este género de dificultades. Carnap, Hempel y otros autores pronto se percataron de la imposibilidad, o al menos inverosimilitud, de definir otros tipos de términos, aún más centrales en muchas disciplinas científicas, en función de términos observacionales en el sentido fisicalista (y por supuesto también en el fenomenalista). Aparte de la categoría de los conceptos disposicionales, hay por lo menos otros tres grandes tipos de términos teóricos para los que, por razones análogas, no parece posible una reducción definicional a  $L_0$ . Estos tipos son:

a) *Conceptos métricos o magnitudes*

Son la inmensa mayoría de los conceptos fundamentales de la física y de buena parte de las demás disciplinas (ejemplos: «longitud», «masa», «energía», «carga eléctrica», «vida media», «cociente de inteligencia», «precio de equilibrio»). Estos conceptos son funciones en el sentido matemático; asignan números reales a objetos empíricos. La razón por la que estos conceptos no son definibles en términos observacionales no es tanto que, por ser funciones reales, pueden tomar valores irracionales que en consecuencia son imposibles de «observar» (nadie puede percibir una longitud de exactamente  $\sqrt{2}$  centímetros), como al principio creyeron Carnap y Hempel<sup>1</sup>; la razón fundamental es más bien que lo característico de las magnitudes físicas (o de otras disciplinas metrizadas) es que haya para ellas diversos procedimientos de medición, asociados a diversas operaciones de laboratorio, y de modo tal que los procedimientos operacionales de medición pueden cambiar drásticamente con el desarrollo científico; ello no obsta para que consideremos que esa serie abierta de procedimientos «observacionales» de medición fijan una y la misma magnitud —precisamente porque esa magnitud forma parte de una y la misma teoría<sup>2</sup>.

1. Esa sería en realidad una razón espuria para la no-observacionalidad de las magnitudes; en efecto, sus valores irracionales podrían definirse, mediante el procedimiento de las sucesiones de Cauchy, a partir de valores racionales, todos los cuales, al menos en principio, son observables (cf. Stegmüller, 1970, 272ss).

2. Para más detalles sobre este aspecto de las magnitudes cf. Moulines, 1986.

b) *Idealizaciones o conceptos ficcionales*

En las ciencias avanzadas, sobre todo cuando están matematizadas, suelen aparecer conceptos de los que la disciplina misma afirma que ninguna entidad real cae bajo ellos. Ejemplos: «punto-masa», «*perpetuum mobile*», «gas ideal», «agente económico». Nótese que no se trata aquí de términos puramente matemáticos (aun cuando su introducción suele ir de la mano de la matematización de una teoría); por el contrario, aunque su referente es vacío, van asociados como idealizaciones o «aproximaciones» a entidades reales; por ejemplo, a partículas reales, máquinas reales, gases reales o seres humanos de carne y hueso. Por otro lado, sería erróneo considerarlos como términos superfluos. Muchas veces (como en los ejemplos considerados), la teoría a la que pertenecen, y que es una buena teoría empírica, no sería ni siquiera formulable sin ellos.

c) *Términos con referente real, pero inobservables por principio*

Son típicos de las teorías microfísicas («fotón», «electrón», *quark*), aunque posiblemente no sólo aparezcan en ellas. (En psicología podemos encontrarnos con tales conceptos también, por ejemplo, el de subconsciente.) En tales casos, no se trata de ficciones o idealizaciones como en el caso precedente; la teoría correspondiente supone que los referentes de esos términos existen en la realidad física (o psíquica); pero a la vez afirma la imposibilidad de observarlos, incluso en un sentido lato de «observar», debido a principios básicos de la propia teoría (el principio de incertidumbre de Heisenberg, por ejemplo, en el caso de «electrón»). Lo que se observa en todos estos casos no es el referente mismo del término sino sus efectos muy indirectos, que la teoría nos permite interpretar como asociados a ellos (por ejemplo, ciertas trazas en una cámara de burbujas). Por ello, a los referentes de los términos del tipo *c*) (y sólo a ellos) podemos caracterizarlos de «entidades teóricas».

## III. ELIMINABILIDAD NO-DEFINICIONAL DE LOS TERMINOS TEORICOS

La inviabilidad del programa definicional para los términos teóricos no implica por sí misma que no sea posible eliminar de alguna otra forma los términos teóricos de las teorías científicas sin perjuicio del contenido empírico de las mismas. Pueden idearse métodos de sustitución de una versión con términos teóricos por otra sin ellos que sea equivalente a la primera desde el punto de vista empírico. Existen de hecho dos métodos de sustitución semejantes, debidos a Craig y Ramsey, respectivamente. El segundo ha tenido mayor influencia en la discusión posterior que el primero.

William Craig demostró (cf. Craig, 1956) con la ayuda de un teorema de la lógica pura previamente demostrado por él mismo (el llamado

«teorema de la interpolación») que una teoría axiomatizada que contiene términos teóricos básicos puede ser sustituida por otra teoría, también axiomatizable, que carezca totalmente de ellos pero que tenga el mismo contenido empírico que la primera en el sentido de que todas y sólo las consecuencias observacionales de la primera lo son también de la segunda. La prueba del teorema de Craig es bastante técnica y hace uso del método llamado de «gödelización» (asignación de números a las expresiones de un lenguaje). Como cuestión de principio puede decirse que, gracias al teorema de Craig, los términos teóricos de cualquier teoría axiomatizada son superfluos. Sin embargo, como cuestión de hecho no lo son, pues la teoría sustitutiva puramente observacional demuestra ser inmanejable: aunque es en principio axiomatizable, contiene un número potencialmente infinito de axiomas pues, dicho brevemente, a cada consecuencia observacional de la teoría original le corresponde su propio axioma.

Más interesante desde un punto de vista epistemológico es el método Ramsey. En este caso, la teoría sustitutiva tiene una estructura algo más manejable que en el caso de Craig. La idea es la siguiente. Sea una teoría empírica axiomatizada  $T$  cuyos términos teóricos son  $t_1, \dots, t_n$  y los observacionales  $o_1, \dots, o_m$ . Reunamos todos los axiomas de  $T$  en una sola conjunción que simbolizaremos por  $T(t_1, \dots, t_n, o_1, \dots, o_m)$ . Sustituyamos ahora los términos  $t_1, \dots, t_n$  por variables libres  $x_1, \dots, x_n$  que ligaremos mediante un cuantificador existencial, obteniendo el enunciado:

$$\exists x_1, \dots, x_n T(x_1, \dots, x_n, o_1, \dots, o_1, \dots, o_m).$$

(Éste es el llamado enunciado Ramsey de la teoría  $T$ .) Lo que Ramsey mostró —la prueba es bastante sencilla— es que las consecuencias observacionales del enunciado Ramsey de  $T$  son exactamente las mismas que las de la propia  $T$ . Dicho intuitivamente, si sólo nos interesáramos por el contenido observacional de una teoría, podríamos hacer como si sus términos teóricos no designaran nada, sino que fueran sólo variables que podemos relacionar con los términos observacionales (que sí tienen significado preciso) y cuantificar existencialmente, de tal modo que el enunciado complejo resultante tendrá las mismas consecuencias observacionales que la versión original de la teoría.

No hay que entender tampoco el método Ramsey como una recomendación práctica de «reescritura» de las teorías científicas, pero sí como la «prueba» de que, desde un punto de vista *epistemológico*, los términos teóricos son superfluos; siempre y cuando, claro está, concordemos en que lo único que nos interesa conocer de la realidad son sus aspectos observables.

A un nivel más profundo, lo que el método Ramsey indica es que, en realidad, los términos teóricos son expresiones sin significado fijo, de «interpretación abierta». Y de ahí, a su vez, se desprende que los enunciados teóricos aislados no son verdaderos enunciados en el sentido de que por

sí solos puedan ser verdaderos o falsos, sino que son más bien formas proposicionales. Sólo el enunciado Ramsey, como unidad global significativa de una teoría, es verdadero o falso.

Es importante notar que el método Ramsey es independiente de cómo se trace la demarcación entre los términos teóricos y los observacionales, y en especial de si esa distinción se hace de manera universal y absoluta, o bien sólo relativa a cada teoría considerada. Cuando presuponemos que la distinción entre términos teóricos y observacionales no es dependiente de cada teoría, es decir, cuando presuponemos un criterio universal y absoluto de teoriedad, nos hallamos en el marco de lo que se ha llamado «la concepción de los dos niveles (del lenguaje científico)».

#### IV. LA CONCEPCION DE LOS DOS NIVELES

Esta concepción es un componente importante de la filosofía clásica o estándar de la ciencia, aunque por supuesto no es su único aspecto. Si adoptamos un criterio universal de teoriedad, aplicable por igual a cualquier teoría de cualquier disciplina científica, la imagen que resulta de la «estructura profunda» conceptual de la ciencia es la de un inmenso edificio de dos pisos (el superior mucho «más alto» que el inferior): el piso teórico y el observacional. Cada teoría científica viene representada, por así decir, por una sección vertical del edificio, conectando ciertas habitaciones superiores con ciertas habitaciones inferiores. Los enunciados puramente teóricos (los que se hallan sólo en el piso superior) nos permiten sacar consecuencias puramente observacionales (que se hallan en el piso inferior) con la ayuda de enunciados mixtos teórico-observacionales (una especie de ascensores entre uno y otro piso), las llamadas «reglas de correspondencia». Por el método Ramsey sabemos que, si sólo nos interesa averiguar lo que ocurre en el piso inferior, podemos prescindir del significado concreto de los enunciados del piso superior y tomarlos sólo como herramientas prácticas para pasar de una estancia del piso inferior a otra estancia del mismo piso inferior (explicaciones y predicciones observacionales).

Esta concepción no sólo atañe a la relación entre conceptos observacionales y teóricos, sino a la idea misma de teoría científica. Ya en una de las primeras expresiones clásicas de esta concepción leemos:

Una teoría científica puede compararse, por tanto, a una red espacial compleja: sus términos vienen representados por los nudos, mientras que los hilos que los conectan corresponden, en parte, a las definiciones y, en parte, a las hipótesis fundamentales y derivadas incluidas en la teoría. El sistema en su conjunto flota, por así decir, por encima del plano de la observación y está anclado en él por las reglas de interpretación (= reglas de correspondencia). Éstas pueden concebirse como cuerdas que no son parte de la red pero que vinculan ciertas partes de la misma con lugares específicos en el plano de la observación (cf. Hempel, 1952).

## V. CRITICAS A LA CONCEPCION DE LOS DOS NIVELES

La distinción absoluta y universal entre lo teórico y lo observacional, y la concepción de las teorías concomitantes, han sido sometidas a crítica radical por diversos autores, tanto por lo que se refiere al concepto de observacionalidad como al de teoriedad. Los filósofos «historicistas» de la ciencia, como Norwood R. Hanson, Thomas S. Kuhn y Paul K. Feysabend, han sostenido la llamada «tesis de la carga teórica universal», según la cual todo concepto científico (e incluso los de la vida cotidiana) están impregnados de teorías implícitas por lo que, en definitiva, todo concepto es teórico. Por tanto es espuria cualquier división entre dos niveles conceptuales en la ciencia.

El alcance de la tesis de la carga teórica universal posiblemente ha sido sobrevalorado en la discusión contemporánea. Es plausible admitir que los conceptos observacionales están, en algún sentido, «impregnados» de supuestos teóricos (aunque habría que precisar qué significa ello exactamente); pero de ahí a concluir que no puede trazarse ninguna distinción con sentido, en ningún contexto, entre nociones observacionales y teóricas, hay un salto lógico. Por lo demás, la tesis de la carga teórica universal tomada literalmente (como parece proponerla Feysabend) conduce inevitablemente a un subjetivismo radical y a una auto-refutación (cf. Scheffler, 1967, c. 1 y 3). El único aspecto realmente plausible de la tesis, que se desprende de la mayoría de ejemplos históricos aducidos por Hanson y otros, es que toda observación científica presupone una asunción previa de un sistema conceptual clasificatorio; pero calificar tal sistema de «teoría» proviene de un uso inflacionario del término (cf. Nola, 1987).

Más pertinente es la crítica, expresada entre otros por el propio Hempel (cf. Hempel, 1971) de que el concepto de observacionalidad es históricamente relativo: lo que para los científicos de una época pasaba por ser un constructo teórico, en una época posterior (gracias al entreno y al desarrollo tecnológico) puede convertirse en observacional (cf. también Scheffler, *op. cit.*).

Otro tipo de críticas son las que atañen a la idea de concepto teórico, a su insuficiente o errónea elucidación. Desde diversas perspectivas se ha señalado que entender los términos teóricos de manera «negativa», simplemente como aquellos que no son observacionales, es inadecuado. Así, Bar-Hillel aduce que no hay que confundir la dicotomía observacional/no-observacional con la dicotomía teórico/no-teórico, y que la que realmente interesa para el análisis lógico de las teorías científicas es la segunda (cf. Bar-Hillel, 1970). En una línea análoga hay que ver lo que Stegmüller en su momento popularizó como «el reto de Putnam» (cf. Stegmüller, 1973, 51). En efecto, Putnam criticaba en un artículo de 1962 el modo como se había llevado a cabo hasta entonces la discusión en torno a los términos teóricos, al no poner en claro antes que nada qué es lo realmente *distintivo* de tales términos y cuál es la relación que

guardan con las teorías de las que proceden (cf. Putnam, 1962, 243). Para Putnam, la problemática estaba obviamente conectada con una concepción (hasta entonces errónea) de lo que son las teorías.

El «reto de Putnam» fue desoído por la mayoría de filósofos de la ciencia contemporáneos, excepto por Sneed y Stegmüller. Estos autores se propondrían darle una respuesta precisa y adecuada, y con ello iniciarían una nueva concepción de las teorías científicas, que más tarde se ha dado en llamar «la concepción estructuralista».

#### VI. EL CRITERIO DE *T*-TEORICIDAD EN LA CONCEPCION ESTRUCTURALISTA

El nuevo enfoque de los términos teóricos no es la única característica original de la concepción estructuralista. En realidad, ésta es una meta-teoría de las teorías empíricas relativamente compleja (al menos más compleja que la concepción de los dos niveles), que intenta apresar, mediante un aparato modelo-teórico refinado, diversos aspectos esenciales de las teorías empíricas y sus relaciones mutuas. La noción de «término *T*-teórico» juega en la metateoría estructuralista un papel ciertamente importante, pero junto a otras nociones y distinciones igualmente fundamentales. Aquí no podemos entrar, ni siquiera someramente, en los principios básicos de esa concepción<sup>3</sup>. A los efectos de la discusión presente baste indicar que la idea básica de la concepción estructuralista consiste en tomar modelos (en tanto que estructuras) y no enunciados como las unidades fundamentales del conocimiento científico. Cada teoría viene caracterizada por un conjunto de modelos de estructura idéntica, o dicho más exactamente, por un conjunto de modelos *potenciales* y un conjunto de modelos *actuales*, subconjunto del primero. El primer conjunto corresponde a lo que podríamos llamar el «marco conceptual» de la teoría; el segundo añade a ese marco las leyes propias con contenido empírico.

Cada modelo (potencial o actual) es una estructura que consta de uno o varios dominios de objetos más una serie de relaciones y/o funciones. En general, estas relaciones y funciones serán, según el estructuralismo, de dos tipos: *T*-teóricas o *T*-no-teóricas (siendo *T* la teoría caracterizada por dichos modelos). Es decir, se prescinde de la noción de observacionalidad (que se considera ajena a la estructura de las teorías científicas aunque no necesariamente irrelevante en otros contextos), y la dicotomía entre términos teóricos se considera no universal como en la concepción de los dos niveles, sino relativa a cada teoría dada: los términos *T*-no-teóricos vienen fijados por medios externos a esa teoría *T* y por consiguiente representan su base de contrastación, mientras que los

3. El compendio sistemático más completo hasta ahora es: Balzer/Moulines/Sned, 1987. Pueden encontrarse exposiciones más introductorias, en lengua castellana, en Moulines, 1982; Rivadulla, 1984 y Echeverría, 1989.

términos  $T$ -teóricos vienen fijados por las propias leyes de  $T$  y determinan la capacidad explicativa y predictiva de la teoría.

El criterio original propuesto por Sneed (cf. Sneed, 1971, c. II) se refería sólo a las funciones métricas y su formulación no era todo lo precisa que sería de desear. Rezaba así: «Una función [métrica] es  $T$ -teórica si todos los métodos de medición de la misma presuponen la aplicabilidad de las leyes de la teoría». Tanto la noción de «método de medición» como la de presuposición quedaban aquí envueltas en cierta vaguedad.

Balzer y Moulines propusieron más tarde una generalización y precisión modelo-teórica de la idea intuitiva original de Sneed (cf. Balzer, Moulines, 1980). El criterio de  $T$ -teoricidad se formula para cualquier tipo de conceptos empíricos, sean métricos o no. Se introduce primero formalmente la noción de método de determinación (de un concepto dado) como una clase de modelos potenciales de la teoría en cuestión que cumplen ciertas condiciones; luego se define como  $T$ -teórico cualquier término de  $T$ , todos cuyos métodos de determinación no sólo son modelos potenciales, sino además *actuales* de  $T$  (o sea, cumplen las leyes de  $T$ ). Los ejemplos de aplicación del criterio que dan Balzer y Moulines muestran que la nueva formulación es intuitivamente adecuada y más precisa que la original. Sin embargo, tiene aún algunas limitaciones y sobre todo un componente pragmático ineludible. Posteriormente se han hecho otros intentos, en una línea parecida, de formalizar aún mejor el criterio de  $T$ -teoricidad (cf. Gähde, 1987; Forge, 1984; Balzer, 1985), pero su exposición requeriría de detalles técnicos que romperían el marco del presente artículo.

#### BIBLIOGRAFIA

- Ayer, A. J. (comp.) (1959), *Logical Positivism*, Free Press, Glencoe, Ill, v.e. Schmill, U. et al., *El positivismo lógico*, FCE, México, 1965.
- Balzer, W. (1985), «A New Definition of Theoreticity»: *Dialectica* 39, 127-145.
- Balzer, W., Moulines, C.U., (1980), «On Theoreticity»: *Synthese* 44, 467-494.
- Balzer, W., Moulines, C.U., Sneed, J.D. (1987), *An Architectonic for Science*, Reidel, Dordrecht.
- Bar-Hillel, Y. (1970), «Neo-Realism vs. Neopositivism: A Neo-Pseudo Issue», en Bar-Hillel, Y., *Aspects of Language*, Magnes Press, Jerusalén.
- Carnap, R. (1928), *Der logische Aufbau der Welt*, Felix Meiner, Hamburg, 2ª edic. (1961), v.e. Mues, L., *La construcción lógica del mundo*, UNAM, México, 1988.
- Carnap, R. (1936-37), «Testability and **Meaning**»: *Philosophy of Science* 3 y 4.
- Craig, W. (1956), «Replacement of Auxiliary Expressions»: *Philosophical Review* LXV.
- Echeverría, J. (1989), *Introducción a la metodología de la ciencia*, Barcanova, Barcelona.
- Forge, J. (1984), «Theoretical Functions in Physical Science»: *Erkenntnis* 21, 1-29.
- Gähde, U. (1983), *T-Theoretizität und Holismus*, Peter Lang, Frankfurt a. M., Bern.
- Hempel, C. G. (1952), «Fundamentals of Concept Formation in Empirical Science», en Neurath, O., Carnap, R., Morris, Ch., *International Encyclopedia of Unified Science* (t. 2-7), University of Chicago Press, Chicago.
- Hempel, C. G. (1971), «The Meaning of Theoretical Terms: A Critique of the Standard Empiricist Construal», en Suppes, P. (comp.), *Logic, Methodology and Philosophy of*

- Science. IV. Proceedings of the 1971 International Congress*, North-Holland, Amsterdam.
- Mach, E. (1868), «Über die Definition der Masse»  
für *physikalische Technik, für mathematische und astronomische Instrumentenkunde 4*.
- Mach, E. (1883), *Die Analyse der Empfindungen und das Verhältnis des Physischen zum Psychischen*, 6a edic. (1911), v.e. Ovejero, E., *Análisis de las sensaciones*. Daniel Jorro, Madrid, 1925.
- Moulines, C.U. (1975), «La génesis del positivismo en su contexto científico»: *Dianoia*, 31-49.
- Moulines, C. U. (1976), «Los fundamentos de la filosofía natural de Isaac Newton»: *noia*, 27-43.
- Moulines, C. U. (1982), *Exploraciones metacientíficas*, Alianza, Madrid.
- Moulines, C. U. (1986), «The Ways of Holism», *Nous* XX, 3, 313-330.
- Moulines, C. U. (1991), «Making Sense of Carnap's "Aufbau"»: *Erkenntnis* 35, 263-286.
- Nola, R. (1987), «Some problems concerning the theory-ladenness of observation»: *Dialectica* 41, 273-292.
- Putnam, H. (1962), «What Theories Are Not», en Nagel, E., Suppes, P., Tarski, A. (comps.), *Logic, Methodology and Philosophy of Science*, Stanford.
- Rivadulla, A. (1984), *Filosofía actual de la ciencia*, Alianza, Madrid.
- Scheffler, I. (1967), *Science and Subjectivity*, Bobbs-Merrill, Indianapolis.
- Sneed, J. D. (1971), *The Logical Structure of Mathematical Physics*, Reidel, Dordrecht.
- Stegmüller, W. (1970), *Theorie und Erfahrung*, Springer, Berlin, Heidelberg, v.e. Moulines, C.U., *Teoría y experiencia*, Ariel, Barcelona, 1979.
- Stegmüller, W. (1973), *Theorienstrukturen und Theorien-Dynamik*, Springer, Berlin, Heidelberg, v.e. Moulines, C.U., *Estructura y dinámica de teorías*, Ariel, Barcelona, 1983.
- Suppes, P. (1957), *Introduction to Logic*, Van Nostrand, New York.