

**PUTNAM, H. (1988:1).** *Razón, verdad e historia*. Madrid: Tecnos.

El propósito que anima la presente obra es acabar con la presión asfixiante que unas cuantas dicotomías parecen ejercer tanto sobre el pensamiento de los filósofos como sobre el de legos. La principal de éstas es la dicotomía entre concepciones objetivas de la verdad y de la razón.

... una vez que una dicotomía como la existente entre lo «objetivo» y lo «subjetivo» se convierte en una dicotomía aceptada, y aceptada no como mero par de categorías, sino como una caracterización de tipos de concepciones y estilos de pensamiento, los pensadores comienzan a ver los términos de la dicotomía casi como etiquetas ideológicas. Muchos filósofos (quizá la mayoría) sostienen hoy alguna versión de la **teoría de la verdad-copia, concepción de acuerdo con la cual un enunciado es verdadero sólo en el caso de que se «corresponda con los hechos (independientes de la mente)»** ... consideran que la única alternativa a ésta es negar la objetividad de la verdad y capitular con la idea de que todos los esquemas de pensamiento y todos los puntos de vista son desesperadamente subjetivos... una audaz minoría (Kuhn... Foucault) se alinean bajo la etiqueta opuesta. Estos últimos están de acuerdo con que la alternativa a la concepción ingenua de la verdad-copia es considerar subjetivos a los sistemas de pensamiento, a la ideología, e incluso (en el caso de Kuhn y Feyerabend a las teorías científicas y pasan a proponer con vigor una perspectiva realista y subjetiva.

**PUTNAM, H. (1988:2).** *Razón, verdad e historia*. Madrid: Tecnos.

La concepción que voy a defender juzga, dicho sin rodeos, que **hay una relación sumamente estrecha entre las nociones de *verdad* y de racionalidad**; expresémoslo más directamente si cabe: que el único criterio para decidir lo que constituye un hecho es lo que es racional aceptar. (Con esto quiero afirmar algo completamente literal y sin excepción, de forma que si admitimos como posible el que sea racional aceptar que una pintura es bella, entonces es posible que sea un hecho que la pintura es bella). Mi concepción admite hechos de valor. Pero la relación entre la aceptabilidad racional y a verdad se da entre dos nociones distintas. **Un enunciado puede ser racionalmente aceptable en un tiempo y no ser verdadero. En este trabajo trataré de preservar esta intuición realista.**

Sin embargo, **no creo que la racionalidad se defina mediante un conjunto de «cánones» o «principios» invariables; los principios metodológicos están relacionados con nuestra visión del mundo, incluyendo la visión que tenemos de nosotros mismos como una parte del mundo, y varían con el tiempo.** De modo que estoy de acuerdo con los filósofos subjetivistas en que no hay ningún *órgano fijo* y ahistórico que defina lo que es racional... no concluyo que la razón pueda ser (o evolucionar hacia) algo, ni caigo en una fantástica mezcla de relativismo cultural y «estructuralismo», como hacen los filósofos franceses. **Considero que la dicotomía: cánones de racionalidad ahistóricos e invariantes o relativismo cultural, está anticuada.**

**PUTNAM, H. (1988:2-3).** *Razón, verdad e historia*. Madrid: Tecnos.

Otro rasgo de mi concepción es que no confina la racionalidad en el laboratorio científico, ni la considera fundamentalmente diferente dentro y fuera de éste. La idea contraria me parece una resaca del positivismo; de la idea de que el mundo científico está configurado en cierta manera por «sense-data» y de la idea de que los términos de las ciencias de laboratorio están definidos operacionalmente. .. Pero la idea empirista de que los «sense-data» constituyen una especie de «planta baja» al menos para una parte de nuestro conocimiento, será reexaminada a la luz de los que tenemos que decir a propósito de la verdad y de la racionalidad.

**En resumidas cuentas, presentaré una perspectiva según la cual la mente no «copia» simplemente un mundo** que sólo admita la adscripción de La Teoría Verdadera. Pero desde mi punto de vista, **la mente no constituye el mundo** (ni siquiera en cuanto que estando sujeta a la constrictión impuesta por «cánones metodológicos» y «sense-data» independientes de la mente).

PUTNAM, H. (1988:**1-3**). *Razón, verdad e historia*. Madrid: Tecnos.

Una de las finalidades de mi estudio acerca de la racionalidad es ésta: tratar de mostrar que **nuestra noción de racionalidad es, en el fondo, solamente una parte de nuestra concepción del florecimiento humano, es decir, de nuestra idea de lo bueno**. En el fondo, la verdad depende de lo que recientemente se ha denominado «valores». Y lo que afirmamos anteriormente con respecto a la racionalidad y a la historia, se aplica también al valor y a la historia; **no hay ningún conjunto de «principios morales», ahistórico, que defina de una vez por todas en qué consiste el florecimiento humano; pero esto no significa que todo sea meramente cultural y relativo**. Los puntos de vista al uso de la verdad están enajenados; causan que se pierda una u otra parte de uno mismo y del mundo, y que se conciba al mundo consistiendo simplemente en partículas elementales desplazándose caóticamente en el vacío (la visión fisicalista, que considera que la descripción científica converge hacia la Teoría de la Verdad) o en «sense-data actuales o posibles» (la teoría empirista más rancia) o a **negar que haya un mundo** como algo opuesto a un puñado de historias que fabricamos por diversas razones (principalmente inconscientes). Mi propósito en esta obra es esbozar las ideas directrices de una perspectiva no enajenada.

**PUTNAM, H. (1988:5).** *Razón, verdad e historia*. Madrid: Tecnos.

### **Cerebros en una cubeta**

Una hormiga se arrastra lentamente sobre la arena. Conforme avanza va trazando en ésta una línea. Pero por azar, la línea se desvía y vuelve sobre sí misma, de tal forma que acaba pareciendo una reconocible caricatura de Winston Churchill. ¿Ha trazado la hormiga un retrato de Winston Churchill?

La mayoría de la gente, tras reflexionar un poco, contestaría que no. Después de todo, la hormiga nunca ha visto a Churchill, ni siquiera un retrato suyo, ni tampoco tenía intención de representarlo. Simplemente trazó una línea (y ni siquiera este acto fue intencional), la línea que nosotros podemos ver como un retrato de Winston Churchill.

Podemos expresar esto afirmando que la línea no representa por sí misma. La semejanza ... con las facciones de Winston Churchill no es una condición suficiente para que algo represente o se refiera a Churchill. ... Si la semejanza no es condición necesaria ni suficiente para que alguna cosa represente a otra, ¿cómo demonios puede una cosa representar (o estar en lugar de. Etc.) otra diferente?

... Por otra parte, supongamos que la línea tiene forma WINSTON CHURCHILL, y que este hecho es un mero accidente ... Entonces los «caracteres impresos» WINSTON CHURCHILL no habrían representado a Winston Churchill, a pesar de que sí lo hacen cuando aparecen hoy en casi todos los libros.

De forma que puede antojársenos que lo que se necesita para la **representación, o lo que se necesita principalmente para la representación, es la intención.**

**PUTNAM, H.** (1988:6). *Razón, verdad e historia*. Madrid: Tecnos.

Pero para tener la intención de que algo, siquiera el lenguaje privado (incluso las palabras «Winston Churchill» repetidas mentalmente y no oídas) represente a Churchill, debo ser capaz de pensar en Churchill, para empezar. Si las líneas en la arena, los ruidos, etc., no pueden representar nada «en sí mismos», entonces ¿cómo es que pueden hacerlo las formas del pensamiento? ¿Cómo puede el pensamiento alcanzar y «aprehender» lo que es externo?

Algunos filósofos han dado el salto desde estas reflexiones hasta lo que ellos consideran como una prueba de la naturaleza no física de la mente. El argumento es simple; **lo que dijimos acerca de la curva de la hormiga también aplica a cualquier objeto físico. Ningún objeto físico tiene por sí mismo la capacidad de referirse a una cosa más bien que a otra; no obstante, es obvio que los pensamientos de la mente si lo logran. De modo que los pensamientos (y por ende, la mente) poseen una naturaleza esencialmente distinta de la de los objetos físicos. Tienen la característica distintiva de la *intencionalidad* –pueden referirse a otras cosas; ningún objeto físico tiene «intencionalidad», salvo la intencionalidad que se deriva de su uso por parte de la mente. O eso se pretende. Pero esto es ir demasiado deprisa; postular misteriosos poderes mentales no resuelve nada. A pesar de todo el problema es real. ¿Cómo es posible la intencionalidad? ¿Cómo es posible la referencia?**

**PUTNAM, H.** (1988:6-7). *Razón, verdad e historia*. Madrid: Tecnos.

## **TEORÍAS MÁGICAS DE LA REFERENCIA**

.. Ciertos pueblos primitivos creen que algunas representaciones (en particular los nombres) tienen una conexión necesaria con sus portadores; creen que saber el «verdadero nombre» de alguien o algo les otorga poder sobre ese alguien o algo. Este poder procede de la *conexión mágica* entre el nombre y el portador del nombre; pero una vez que nos percatamos de que el nombre sólo tiene una conexión contextual, contingente y convencional con su portador, es difícil ver por qué el conocimiento del nombre ha de tener alguna significación mística.

Es importante darse cuenta de que las *imágenes mentales*, y, en general, las representaciones mentales, les ocurre lo mismo que a los dibujos físicos; la conexión que tienen las representaciones mentales con lo que representa no es más necesaria que la que tienen representaciones físicas. La suposición contraria es un vestigio del pensamiento mágico.

**PUTNAM, H.** (1988:6-7). *Razón, verdad e historia*. Madrid: Tecnos.

El problema quizá se capte más fácilmente en el caso de las imágenes mentales (quizá el primer filósofo que captó la enorme importancia de este problema... fue Wittgenstein).

Supongamos que en alguna parte existe un planeta en el cual se han desarrollados seres humanos ... Supongamos que esos humanos, si bien son como nosotros, nunca han visto un *árbol*. Supongamos que nunca se han imaginado un árbol (la vida vegetal que existe en su planeta son los líquenes). Supongamos que cierto día, una nave que pasa por su planeta sin establecer contacto con ellos, arroja sobre éste el dibujo de un árbol. Imaginémosles devanándose los sesos ante el dibujo. ¿Qué demonios es eso? Se les ocurre toda clase de especulaciones: un edificio, un baldaquín, e incluso alguna especie de animal. Pero supongamos que ni siquiera se aproximan a saber de qué se trata.

Para nosotros la pintura es la representación de un árbol. Para aquellos humanos el dibujo únicamente representa un objeto extraño, de naturaleza y función desconocidas. Supongamos que, como resultado de ver el dibujo, uno de ellos tiene una **imagen mental que es exactamente como mis imágenes mentales de los árboles**. Su imagen mental no es la representación de un árbol. Sólo es la representación del extraño objeto (el que sea) que representa la misteriosa pintura.

Pese a esto, alguien podría argumentar que la imagen **mental es de hecho la representación de un árbol**, ya que, en primer lugar, el dibujo que provocó tal imagen mental era la representación de un árbol. **Hay una cadena causal** desde los árboles reales hasta la imagen mental, aun cuando esta sea muy extraña.

Podemos imaginar que la nave espacial que arrojó el «dibujo» procedía de un planeta en el que no se sabía nada sobre los árboles. En tal caso, pese a que los seres humanos tendrían imágenes cualitativamente idénticas a mi imagen de árbol, esas imágenes no representan más a un árbol que a cualquier otra cosa arbitraria.



**Lo mismo ocurre con las palabras.** Un discurso impreso podría parecer una descripción de un árbol, pero si fueron los monos quienes lo produjeron golpeando fortuitamente las teclas de una máquina de escribir durante millones de años, entonces las palabras de ese discurso no se refieren a nada. Si alguien las memorizase y las repitiese mentalmente sin entenderlas, entonces cuando fuesen pensadas tampoco se referirán a nada.

Imaginemos que la persona que está repitiendo mentalmente estas palabras ha sido hipnotizada. Supongamos que tales palabras están en japonés, y que el hipnotizado se le ha dicho que entiende ese idioma, Supongamos que cuando piensa esas palabras experimenta algo así como un «sentimiento de comprensión». (Aunque alguien irrumpiese en su flujo mental y le preguntase qué significan las palabras que está pensando, descubriría que no podría decirlo). Quizá la ilusión sea tan perfecta que incluso podrá engañar a un telépata japonés. Pero si no es capaz de emplear las palabras en los contextos apropiados ni de responder preguntas con respecto a los que «piensa», entonces no las entendió.

Combinando estas historias de ciencia-ficción, podemos idear un caso en el que alguien piensa ciertas palabras que constituyen de hecho la descripción de un árbol en algún lenguaje, y simultáneamente tiene unas imágenes mentales apropiadas, pero ni comprende las palabras ni sabe lo que es un árbol. Incluso podemos imaginar que las imágenes mentales fueron provocadas por un derrame de pintura (aunque la persona ha sido hipnotizada e inducida a pensar que son imágenes de algo apropiado a su pensamiento —sólo que, si se le pregunta de qué son imágenes, no podrá responder). Y podemos imaginar que ni el hipnotizador ni el hipnotizado han oído hablar del lenguaje en el que este último está pensando— quizá sea una mera coincidencia el que estas «oraciones sin sentido», tal como las considera el hipnotizador, sean la descripción de un árbol en japonés. En resumidas cuentas, cualquier cosa que pase ante su mente puede ser cualitativamente idéntica a los que estaba pasando por la mente de un hablante japonés que pensaba verdaderamente en árboles—pero ninguna de ellas se refería a árboles.

PUTNAM, H. (1988:**8**). *Razón, verdad e historia*. Madrid: Tecnos.

Todo esto es realmente imposible... a pesar de todo, no es lógicamente imposible, ni siquiera físicamente imposible. Podría suceder (es compatible con las leyes de la física, y quizá también con las condiciones actuales del universo, si hubiese seres inteligentes sobre otros planetas). Y si si sucediese, sería una sorprendente demostración de una importante verdad conceptual: **ni siquiera un amplio y complejo sistema de representaciones verbales y visuales tiene una conexión intrínseca. Mágica, dada de una vez por todas, con lo que representa**; una conexión independiente del modo en que fue causada y de lo que constituyen las disposiciones del sujeto hablante o pensante. Y esto es cierto tanto si el sistema de representaciones (palabras, imágenes, en nuestro ejemplo) está implementado físicamente —las palabras son palabras escritas o habladas y los dibujos son dibujos físicos— o tan sólo concebido mentalmente. **Ni las palabras del pensamiento ni las imágenes mentales representan intrínsecamente aquello de lo que tratan.**

PUTNAM, H. (1988:**8-9**). *Razón, verdad e historia*. Madrid: Tecnos.

## **EL CASO DE LOS CEREBROS EN UNA CUBETA**

He aquí una posibilidad de ciencia-ficción discutida por los filósofos; imaginemos que un ser humano... ha sido sometido a una operación por un diabólico científico. El cerebro de tal persona (su cerebro, querido lector) ha sido extraído del cuerpo y colocado en una cubeta de nutrientes que lo mantienen vivo. Las terminaciones nerviosas han sido conectadas a una computadora supercientífica que provoca en esa persona la ilusión de que todo es perfectamente normal. Parece haber gente, objetos, cielo., etc. pero en realidad todo lo que la persona (usted) está experimentando es resultado de impulsos electrónicos que se desplazan desde la computadora hasta las terminaciones nerviosas. La computadora es tan ingeniosa que si la persona intenta alzar su mano, el «feedback» que procede de la computadora le provocará que «vea» y «sienta» que su mano está alzándose. Por otra parte, mediante una simple modificación del programa, el diabólico científico puede provocar que la víctima «experimente» (o alucina) cualquier situación o entorno que él desee. También puede borrar la memoria de funcionamiento del cerebro, de modo que la víctima crea que siempre ha estado en ese entorno. La víctima puede creer que está sentado, leyendo estas mismas palabras acerca de la suposición, divertida aunque bastante absurda, de que hay un diabólico científico que extrae cerebros de los cuerpos y los coloca en una cubeta de nutrientes que los mantiene vivos...

PUTNAM, H. (1988:9). *Razón, verdad e historia*. Madrid: Tecnos.

Supongamos esta vez que la máquina automática está programada para ofrecernos a todos una alucinación colectiva, en lugar de unas cuantas alucinaciones separadas y sin relación. De forma que cuando me parece estar hablando con usted, a usted le parece estar oyendo mis palabras. Mis palabras realmente no llegan a sus oídos, por supuesto —porque usted no tiene oídos(reales), ni yo tengo boca o lengua reales— Pero cuando emito mis palabras, lo que ocurre en realidad es que los impulsos aferentes se desplazan desde mi cerebro hasta el ordenador, el cual a su vez provoca que yo «oiga» mis palabras y «sienta» el movimiento de mi lengua, y que usted «oiga» mis palabras, y me «vea» hablando, etc. En este caso, nos comunicamos realmente, hasta cierto punto. Yo no estoy equivocado con respecto a su existencia real (sólo lo estoy con respecto a la existencia de su cuerpo y del «mundo externo», aparte de los cerebros). En cierta medida, tampoco importa que «el mundo entero» sea una alucinación colectiva; después de todo, cuando me dirijo a usted, usted oye realmente mis palabras, si bien el mecanismo no es el que suponemos...

Deseo formular ahora una pregunta que parecerá obvia y bastante estúpida... pero que tal vez nos sumerja con cierta rapidez en auténticas profundidades filosóficas. Supongamos que toda esta historia fuera realmente verdadera. **Si fuéramos cerebros en una cubeta, ¿podríamos decir o pensar que lo somos?**

Voy a argumentar en favor de la respuesta «no, no podemos»... **No puede ser verdadera porque, en cierto modo, se autorrefuta...**

... estaba estudiando el teorema de Skolem-Lówenheim y vi de repente una conexión entre este teorema y algunos argumentos de las *Investigaciones filosóficas* de Wittgenstein.

PUTNAM, H. (1988:**11**). *Razón, verdad e historia*. Madrid: Tecnos.

Un «supuesto que se autorrefuta» es aquel cuya verdad implica su propia falsedad. Por ejemplo, considérese la tesis de que *todos los enunciados generales son falsos*. Este es un enunciado general, de forma que si es verdadero, debe ser falso. Por lo tanto es falso. En ciertas ocasiones decimos que una tesis se «autorrefuta» si la misma suposición de que la tesis es tomada en cuenta o enunciada implica su falsedad. Por ejemplo «no existo» se autorrefuta si *soy yo* (para cualquier «yo») quien lo pienso. De modo que uno puede estar seguro de que existe con sólo pensar en ello (Descartes argumentó).

Demostraré que la suposición de que somos cerebros en una cubeta posee precisamente esta propiedad. Si podemos considerar su verdad o su falsedad, entonces no es verdadera (lo demostraré). Por lo tanto no es verdadera.

La respuesta será (básicamente) ésta: aunque esas personas puedan pensar y «decir» cualquier palabra que nosotros poseemos o digamos, no pueden «referirse» a lo que nosotros nos referimos. En particular, no pueden decir o pensar que son cerebros en una cubeta (incluso pensando «somos cerebros en una cubeta»).

PUTNAM, H. (1988:11). *Razón, verdad e historia*. Madrid: Tecnos.

## EL TEST DE TURING

Supongamos que alguien ha logrado inventar una computadora que pueda mantener realmente una conversación con un sujeto (o con tantas personas como pudiera hacerlo una persona inteligente) ¿Cómo podríamos decidir si la computadora es «*consciente*» o no?

El lógico británico A. Turing propuso en siguiente test: permitamos que un sujeto mantenga una conversación con la computadora y otra conversación con una persona que no conoce. Si no puede decidir quién es la computadora y quién es el humano, entonces (suponiendo que el test se repite en número suficiente de veces con distintos interlocutores) la computadora es consciente. En resumen, una máquina de computar es consciente si puede pasar el «Test de Turing» (las conversaciones no se mantienen cara a cara, por supuesto: el interlocutor no puede conocer la apariencia visual de sus dos contertulios; tampoco debe utilizar la voz, ya que la voz mecánica podría sonar con un matiz distinto al de la voz humana). Imaginemos, más bien, que todas las conversaciones se mantienen con el soporte de una máquina de escribir eléctrica. El interlocutor mecanografía sus enunciados, preguntas, etc., y ambos contertulios —la máquina y el ser humano— responden por medio de un teclado eléctrico. (La máquina puede además *mentir*— si se le pregunta «¿Eres una máquina?» podría replicar «No, soy un asistente de laboratorio»).

Varios autores... han criticado de que este test sea realmente definitivo para la determinación de la conciencia. ...Deseo emplear la idea general del Test de Turing, la idea general de un test *dialogico de competencia*, para un propósito distinto: **explorar la noción de referencia**.

Imaginemos una situación en la que el problema no consiste en determinar si el contertulio es realmente una persona o una máquina, **sino más bien determinar si usa referencialmente las palabras tal y como nosotros lo hacemos**.

PUTNAM, H. (1988:11-12). *Razón, verdad e historia*. Madrid: Tecnos.

El test más obvio consiste, de nuevo, en mantener una conversación y, si no surgen problemas, si el contenido «aprueba» (en el sentido de que no puede ser distinto de alguien de quien sabemos de antemano que habla el mismo lenguaje, se refiere a las mismas clases usuales de objetos, etc.), concluir que se refiere a los objetos tal y como nosotros lo hacemos. Cuando el propósito del Test de Turing sea precisamente el descrito, es decir, determinar la referencia compartida (shared) **lo llamaré *Test de Turing para la referencia***. ... deseo discutir la cuestión de si el recién sugerido Test de Turing para la referencia constituye un **test definitivo para determinar la referencia compartida**.

La respuesta resultará se «NO». El Test de Turing para la Referencia no es definitivo. No hay duda de que es un test excelente en la práctica, pero **no es lógicamente imposible (aunque sin duda altamente improbable) que alguien pase el Test de Turing sin estar refiriéndose a nada**. De aquí se sigue, como veremos, que es posible ampliar nuestra observación **de que las palabras** (y los textos completos de discursos) **no tienen una conexión necesaria con sus referentes**. Aun cuando consideremos las reglas que deciden qué palabras pueden emitirse apropiadamente en ciertos contextos, en lugar de las palabras como tales —aún cuando consideremos programas para el uso de las palabras según la jerga informática—, aún así, **esas palabras no poseen una referencia determinada**, a menos que esos programas se refieran por sí mismos a algo extralingüístico. **Este sería un paso crucial de cara a alcanzar la conclusión de que los cerebros en una cubeta no pueden referirse de ningún modo a nada externo** (y, por consiguiente, no pueden decir que son cerebros en una cubeta).

PUTNAM, H. (1988:**12-13**). *Razón, verdad e historia*. Madrid: Tecnos.

Supongamos, por ejemplo, que me encuentro en una situación ideada por Turing ... y que mi contertulio es efectivamente una máquina. Supongamos que la máquina... está programada para dar bellas respuestas en Castellano en contestación... pero que no tiene órganos sensoriales (salvo la conexión a la máquina de escribir electrónica). (Turing no da por sentado que la posesión de otros órganos sensoriales...). Supongamos que la máquina no sólo carece de ojos y oídos electrónicos, etc., sino que su programa que le permite llevar a cabo el juego de la imitación, también carece de dispositivos para la incorporación de *inputs* desde esos órganos, o para ejercer control sobre su cuerpo. ¿Qué tenemos que decir con respecto a esta máquina?

**Me parece evidente que ni podemos ni debemos atribuir uso referencial a tal ingenio mecánico.** Es cierto que la máquina puede platicar maravillosamente acerca del paisaje de Nueva Inglaterra, por ejemplo. **Pero si tuviese enfrente una manzana, una montaña o una vaca, no podría reconocerlas.**

Tenemos un artilugio para producir oraciones en respuesta a oraciones. Pero esas oraciones no están conectadas con el mundo real. Si *acoplásemos* dos de estas máquinas y les dejásemos disputar el juego de imitación entre sí, continuarían «engañándose» una a otra eternamente, aun cuando el resto del mundo desapareciese. **No existen más razones para considerar que el discurso de la máquina acerca de manzanas se refiere a manzanas del mundo real,** que para considerar que el «retrato» de la hormiga se refiere a Winston Churchill. **Lo que aquí se produce es una ilusión de referencia, significado, inteligencia, etc.,** es el hecho de que tenemos una convención de representación bajo el cual el discurso de la máquina se refiere a manzanas, campanarios, Nueva Inglaterra, etc. De **modo parecido, y por la misma razón, existe la ilusión de que la hormiga ha caricaturizado a Churchill. Sólo que nosotros somos capaces de percibir, manipular y mantener trato con manzanas y campos.**



Existen «reglas de entrada al lenguaje» que nos conducen desde las experiencias con manzanas a proferencias tales como «Veo una manzana» y «reglas de salida del lenguaje» que nos llevan desde decisiones expresadas en forma lingüística («Voy a comprar algunas manzanas») a acciones distintas de la acción de hablar. Como la máquina carece de reglas de entrada o de salida del lenguaje, no hay ninguna razón para considerar el discurso de la máquina ... como algo más que un juego sintáctico... que se parece a un discurso inteligente, a buen seguro. Pero en la misma medida (y no en mayor) que la curva de la hormiga se parece a una mordaz caricatura. ... que la hormiga habría dibujado la misma curva aunque Winston Churchill jamás hubiese existido. Pero no podemos elaborar un argumento que sea completamente paralelo para el caso de la máquina; si las manzanas, los árboles o los campanarios no hubiesen existido, entonces, presumiblemente, los programadores no hubieran dado a luz el mismo programa. A pesar de que la máquina no percibe manzanas, campos o campanarios, sus creadores-diseñadores si lo hacen. **Hay cierta conexión causal entre la máquina y las manzanas del mundo real, etc., por medio de la experiencia perceptiva y el conocimiento de sus creadores diseñadores. Pero esta débil conexión difícilmente puede bastar para la referencia.** No sólo es lógicamente —pese a ser fantásticamente improbable— que la misma máquina hubiese existido sin existir manzanas, campos, campanarios, etc. Aun cuando todas estas cosas cesaran de existir, la máquina seguiría platicando felizmente. **Por este motivo, es absolutamente imposible sostener que la máquina es capaz de referirse a las cosas.**

... el Test de Turin no permite excluir a una máquina programada para no hacer otra cosa *salvo disputar* el juego de la imitación, **y está claro que una máquina con tales características no se refiere a las cosas en mayor medida en que lo hace un tocadiscos.**

**PUTNAM, H.** (1988:14). *Razón, verdad e historia*. Madrid: Tecnos.

... **El hecho de que sean conscientes o inteligentes no significa que sus palabras se refieran a lo que se refieren las nuestras.** La cuestión que nos interesa es: cuando sus verbalizaciones contienen la palabra «árbol», ¿se refiere realmente a árboles? De forma más general: ¿acaso pueden referirse a objetos externos? ...

... **la semejanza cualitativa con algo que representa un objeto (Winston Churchill o un árbol) no hace que una cosa sea por sí misma una representación.** En resumen, cuando los cerebros en una cubeta piensan «hay un árbol delante de mí» no están pensando en árboles reales, ya que no existe nada en virtud de lo cual su pensamiento «árbol» represente árboles reales.

... Si estos cerebros representan, se refieren y piensan acerca de árboles (árboles reales fuera de la cubeta), debe ser gracias al modo en que el programa conecta el sistema lingüístico con *inputs y out-puts* no verbales. Tales inputs existen efectivamente en el mundo de los cerebros en una cubeta (¿de nuevo aquellas terminaciones nerviosas aferentes y eferentes!) pero también veíamos que los «sense-data» producidos por la maquinaria automática no representan árboles (o algo externo) aun cuando parezcan con exactitud a nuestras imágenes de los árboles. Así como una mancha de pintura podría parecer el dibujo de un árbol sin ser el dibujo de un árbol, veíamos que un «dato sensorial» podría ser cualitativamente idéntico a una «imagen de un árbol», sin ser la imagen de un árbol.

PUTNAM, H. (1988:14-15). *Razón, verdad e historia*. Madrid: Tecnos.

**En el caso de los cerebros en una cubeta, ¿cómo es que el hecho de que el lenguaje está conectado mediante el programa con *inputs* sensoriales que no representan intrínseca o extrínsecamente árboles (o cualquier otra cosa externa) puede posibilitar que todo el sistema de representaciones, el lenguaje-en-uso, represente o se refiera a árboles, o a cualquier otra cosa externa?**

**La respuesta es que no puede.** Todo el sistema de *sense-data*, las señales motoras para las terminaciones eferentes y el pensamiento mediatizado verbal o conceptualmente y conectado mediante «reglas de salida del lenguaje» con las señales motoras *outputs*, no tiene más conexiones con árboles que las que la curva de la hormiga tenía con Winston Churchill. Una vez vemos que la semejanza cualitativa (ascendiéndola, si queremos, a identidad cualitativa) entre los pensamientos de los cerebros en una cubeta y los de alguien que exista en el mundo real no implica en modo alguno la mismidad de la referencia, no es difícil ver **que no hay ninguna base para considerar que el cerebro en una cubeta se refiere a cosas externas.**

**PUTNAM, H.** (1988:15-16). *Razón, verdad e historia*. Madrid: Tecnos.

## **LA PREMISA DEL ARGUMENTO**

Según acabamos de afirmar, cuando el cerebro en una cubeta.. piensa «Hay un árbol delante de mí», su pensamiento no se refiere a árboles reales. Según algunas teorías.. podría referirse a árboles que aparecen en la imagen , o a impulsos electrónicos que ocasionan experiencias de árboles. O a las características del programa que son responsables de esos impulsos electrónicos que ocasionan experiencias de árboles, o a las características del programa que son responsables de esos impulsos electrónicos. Nuestras recientes afirmaciones no implican el rechazo de estas teorías, pues **existe una estrecha conexión causal entre el uso de la palabra «árbol» en la cubeta-castellana y la presencia de árboles que aparecen en la imagen...** Según tales teorías, el cerebro *está en lo cierto*, no se equivoca al pensar «Hay un árbol delante de mí». Dada la referencia de «árbol» y de «delante de» en la cubeta-castellana, y suponiendo que una de estas teorías es correcta, entonces las condiciones de verdad de la oración «Hay un árbol delante de mí»... son algo tan simple como el que haya un árbol aparente-en-la-imagen delante de mí en cuestión —en la imagen— o quizá que desde la imagen ha llegado el tipo de impulso electrónico que normalmente produce esta experiencia, o quizá que esté operando el dispositivo de la máquina que produce la experiencia de «árbol delante de mí». Y sin duda alguna, estas condiciones de verdad se satisfacen.

PUTNAM, H. (1988:15-16). *Razón, verdad e historia*. Madrid: Tecnos.

... pero sin duda no se refiere a cubetas reales, ya que el uso de «cubeta» en la cubeta-castellana no tiene conexión causal con cubetas reales (aparte de la conexión que supone el hecho de que **los cerebros en una cubeta no podrían usar la palabra «cubeta» si no estuvieran en presencia de una cubeta particular**... Se sigue que si su «mundo posible» es el mundo real, y somos realmente cerebros en una cubeta entonces lo que queremos decir con «Somos cerebros en una cubeta» es que somos cerebros en una cubeta aparente-en-la-imagen o algo de esta índole... Así, pues, si somos cerebros en una cubeta, entonces la oración «Somos cerebros en una cubeta» afirma algo falso (si es que afirma algo). En resumidas cuentas, si somos cerebros en una cubeta, entonces «Somos cerebros en una cubeta» es falso. Por lo tanto es (necesariamente) falso.

**PUTNAM, H.** (1988:17-17). *Razón, verdad e historia*. Madrid: Tecnos.

**La suposición de que tal posibilidad tenga sentido surge de la combinación de dos errores: (1) tomar demasiado en serio la posibilidad física; y (2) manejar de forma inconsciente una teoría mágica de la referencia, es decir, una teoría en la que ciertas representaciones mentales se refieren necesariamente a ciertas cosas y clases externas. ..**

existe en nuestra cultura (desde el siglo XVIII) la tendencia a considerar la física como nuestra metafísica, es decir, a considerar las ciencias exactas como la descripción del «**verdadero mobiliario último del universo**» por tanto tiempo buscada, la consecuencia inmediata de esa tendencia es también cierta tendencia a considerar la «posibilidad física» como la misma piedra de toque de cómo podrían ser verdadera y realmente las cosas. **La verdad es verdad física, la posibilidad, posibilidad física, y la necesidad, necesidad física, según esta opinión... este punto de vista es erróneo.** La existencia de un «mundo físicamente posible» en el cual somos cerebros en una cubeta (y siempre los fuimos y lo seremos) no significa que posible, verdadera y realmente podríamos ser cerebros en una cubeta. Lo que excluye la posibilidad no es la física, sino la filosofía...

Bien, no hay duda de que mi argumento puede describirse como un argumento conceptual. Pero **cuando la actividad filosófica se describe como una búsqueda de verdades «conceptuales», suena a algo así como una investigación acerca del significado de las palabras. Y, no ha sido ésta nuestra tarea, desde luego.**

PUTNAM, H. (1988:17). *Razón, verdad e historia*. Madrid: Tecnos.

**Nuestra tarea ha consistido en examinar las precondiciones del pensar acerca de algo, representar algo, referirse a algo, etc.** ... Tal procedimiento no es «empírico», pero tampoco es completamente «a priori» ... A pesar de ser falible y de depender de supuestos que podrían ser descritos como «empíricos» (por ejemplo, el supuesto de que la mente no tiene otro acceso a las cosas externas o a las propiedades que el que le proporcionan los sentidos), mi procedimiento se halla en estrecha relación con lo que Kant llamó «investigación trascendental», ya que consiste, repito, en **una investigación sobre las precondiciones de la referencia y, por ende, del pensamiento** —precondiciones que se encuentran incorporadas en la naturaleza de nuestras mentes—, **aunque no son** (como creía Kant) **completamente independientes de suposiciones empíricas.**

**Una de las premisas del argumento es obvia: las teorías mágicas de la referencia son erróneas,** y no sólo erróneas con respecto a las representaciones físicas, sino también en lo que concierne a las mentes. La otra premisa establece la imposibilidad de referirnos a ciertos tipos de cosas, por ejemplo, a *árboles*, sin haber tenido interacción causal con ellas, o en su caso, con otras cosas en cuyos términos puedan describirse las primeras...

**PUTNAM, H.** (1988:17-18). *Razón, verdad e historia*. Madrid: Tecnos.

## **RAZONES PARA NEGAR QUE HAYA CONEXIONES NECESARIAS ENTRE LAS REPRESENTACIONES Y SUS REPRESENTANTES**

Anteriormente apuntaba que algunos filósofos (el más célebre, Bretano) han adscrito una facultad a la mente, **la «intencionalidad»**, que capacita precisamente para *referirse a las cosas*. Evidentemente, descarto que esto constituya una solución. Pero ¿con qué lo hago? ¿He ido demasiado deprisa?

Estos filósofos no afirmaban que fuera posible pensar sin usar ningún tipo de representación. Y aceptarían el argumento ... que comparaba los *sense data* con el «dibujo» de la hormiga... que daba lugar a *sense data* cualitativamente similares a nuestras imágenes visuales de los árboles, **pero no estaban acompañados por ningún concepto de árbol, como muestra de que las imágenes no se refieren a las cosas de un modo necesario. Si existen representaciones mentales que se refieran necesariamente (a cosas externas), deben ser de la naturaleza de los conceptos, no de las imágenes.** Pero, ¿qué son los conceptos?

Cuando practicamos la introspección no percibimos «conceptos» fluyendo por nuestra mente así como así. Dondequiera y cuandoquiera que detengamos el flujo del pensamiento, lo que atrapamos son palabras, imágenes, sensaciones, sentimientos. Cuando digo en voz alta mis pensamientos, no los pienso por segunda vez. Oigo mis palabras como usted las oye. No hay duda de que cuando creo las palabras que profiero me siento en un estado «diferente» a cuando no las creo... Pero puedo imaginar sin dificultad a alguien que piense estas palabras (en el sentido de repetirlas mentalmente) y experimente el mismo sentimiento de comprensión, afirmación, etc., que yo tengo, pero percatándose ... un minuto más tarde de que no entendía lo que acababa de pasar por su mente, y que ni siquiera entendía el lenguaje en el que estaban esas palabras. No pretendo que esto sea muy probable, sino que es perfectamente imaginable.



**PUTNAM, H.** (1988:18-19). *Razón, verdad e historia*. Madrid: Tecnos.

Y esto no muestra que los conceptos son palabras (o imágenes, sensaciones, etc.) sino que atribuir a alguien un concepto o un pensamiento es algo completamente distinto de atribuirle alguna «representación» mental, alguna entidad o evento susceptible de introspección. La razón es decisiva para sostener que los conceptos no sean representaciones mentales que se refieran intrínsecamente a las cosas es que ni tan siquiera son representaciones mentales. **Los conceptos son símbolos que se usan de cierto modo; los símbolos pueden ser públicos o privados, entidades mentales o físicas, pero aun cuando los símbolos sean «mentales» y «privados», el propio símbolo, con independencia del uso, no es el concepto. Y los símbolos no se refieren de por sí intrínsecamente a ninguna cosa.**

Podemos observar esto realizando un experimento mental muy simple. Supongamos que ni usted ni yo podemos distinguir un olmo de un haya. Aún así, afirmamos que la referencia del «olmo» en mi discurso es idéntica a la referencia del «olmo» en el discurso de cualquier otra persona, v.gr. olmos y que la extensión de «haya» es el conjunto de todas las hayas (es decir, el conjunto de cosas de las que la palabra «haya» se predica con verdad) tanto en su discurso como en el mío. **¿Es posible creer que la diferencia entre la referencia del «haya» descansa en una diferencia en nuestros conceptos?** Mi concepto de olmo es exactamente el mismo que mi concepto de haya ... (Esto demuestra de paso, que la determinación de la **referencia es social y no individual**; tanto usted como yo confiamos en los expertos que sí pueden distinguir los olmos de las hayas).

**Contrariamente a lo que afirma la doctrina que se ha mantenido entre nosotros desde el siglo diecisiete, los significados no están en la cabeza.**

**PUTNAM, H.** (1988:19-20). *Razón, verdad e historia*. Madrid: Tecnos.

**los significados no están en la cabeza.**

**Hemos visto que poseer un concepto no consiste en poseer imágenes** (de árboles, por ejemplo— o incluso de imágenes «visuales» o «acústicas» de oraciones, o de discursos enteros, dicho sea de paso), ya que alguien podría poseer cualquier sistema de imágenes que a usted se le antoje sin poseer por ello la capacidad de utilizar las sentencias de modos situacionalmente adecuados, considerando los factores lingüísticos, como los determinantes de la **«adecuación situacional»**. Un hombre podría tener todas las imágenes que usted quiera, y sin embargo, no saber qué hacer cuando se le dice «Señale un árbol», aun cuando haya presente un grupo de árboles. Hasta aquí podría tener la imagen de lo que tiene que hacer, y, con todo, no saber lo que tiene que hacer. Pues la imagen, si no está acompañada por la capacidad de actuar de cierta manera, es sólo imagen, y la capacidad de actuar de acuerdo «con un dibujo» es algo que se puede tener o no. Ese hombre podría representarse a sí mismo señalando un árbol, solamente con el ánimo de considerar algo lógicamente posible, después de que alguien produjese la secuencia de sonidos —que para él no tiene significado— «Por favor, señale un árbol». Y a pesar de todo no sabría que debe señalar un árbol, ni comprendería la exhortación «Señale un árbol».

...los tipos de fenómenos internos que permitimos como posibles expresiones del pensamiento, podrían ofrecer argumentos exactamente similares al anterior que demostrarían que **el entendimiento no está constituido por fenómenos, sino más bien por la capacidad del sujeto en las circunstancias adecuadas.**

**PUTNAM, H.** (1988:20). *Razón, verdad e historia*. Madrid: Tecnos.

... la pretensión de comprender el pensamiento a la luz de la denominada «investigación fenomenológica» se halla fundamentalmente desencaminada; pues lo que los fenomenólogos no ven es que o que están describiendo es la expresión interna del pensamiento, pero la comprensión de esta expresión —la comprensión de los pensamientos de uno mismo— no es algo que acaezca, sino el resultado de una capacidad. Nuestro ejemplo es un hombre que simula pensar en japonés (y engaña al telepata japonés) muestra de inmediato la falibilidad de una aproximación fenomenológica al problema de la comprensión. ...

Por otra parte, consideremos a un hombre (perfectamente posible) que no tiene ningún «monólogo interior». Habla perfectamente castellano, y si le preguntamos cuáles son sus opiniones con respecto a un asunto dado, nos las ofrecerá detenidamente. Mas nunca piensa (con palabras, imágenes, etc.) cuando está hablando en voz alta; ni «pasa nada por su cabeza», salvo que (por supuesto) oye su propia voz hablando, tiene las impresiones sensoriales usuales de sus alrededores, más un «sentimiento general de comprensión». (Quizá tiene el hábito de hablar a solas). Cuando mecanografía una carta, o cuando va de compras, etc. **no experimenta un «flujo de pensamiento» pero sus acciones son inteligentes y propositivas**, y si alguien le aborda y le pregunta «¿Qué está haciendo usted?», ofrecerá una respuesta perfectamente coherente.

Este hombre parece perfectamente imaginable. Nadie vacilaría en decir que es consciente, que no le gusta el rock and roll ... solamente porque no piensa conscientemente excepto cuando habla en voz alta.

**PUTNAM, H.** (1988:20-21). *Razón, verdad e historia*. Madrid: Tecnos.

Dado todo esto se sigue que (a) ningún conjunto de eventos mentales —imágenes o acontecimiento y cualidades más abstractas— constituyen por si mismo la comprensión; y (b) ningún conjunto de eventos mentales es necesario para la comprensión. En particular, **los conceptos no pueden ser idénticos a objetos mentales de ningún tipo**. Pues, aun suponiendo que entendemos por evento mental algo susceptible de introspección, cualquiera que sea éste, **acabamos de ver que podría estar ausente en un hombre que entiende la palabra apropiada (y posee el concepto efectivo) y presente en un hombre que de ningún modo posee el concepto**.

Volviendo ya a nuestra crítica de las teorías mágicas de la referencia... vemos que, por otra parte, estos «objetos mentales» que podemos detectar introspectivamente —palabras, sentimientos, imágenes, etc.— no se refieren a algo intrínsecamente en mayor medida que lo hace el dibujo de la hormiga (y por las mismas razones), mientras que los intentos de postular objetos mentales especiales «conceptos», los cuales sí tienen una conexión necesaria con sus referentes —y que los fenomenólogos entrenados pueden entender— incurrir en un error lógico, porque **los conceptos son (al menos en parte) capacidades, y no cosas que acontecen en la mente. La doctrina que defiende la existencia de representaciones mentales que se refieren a las cosas externas no es sólo mala ciencia natural; es también mala fenomenología y confusión conceptual**.

**PUTNAM, H.** (1988:22). *Razón, verdad e historia*. Madrid: Tecnos.

## **UN PROBLEMA ACERCA DE LA REFERENCIA**

¿Por qué razón nos sorprende que la hipótesis de los cerebros en una cubeta resulte incoherente? La causa es **nuestra tendencia a creer que lo que transcurre en el interior de nuestra cabeza debe determinar lo que queremos decir y aquello a los que se refieren nuestras palabras**. Pero no es difícil atisbar que esta creencia es errónea. Las palabras deícticas ordinarias como yo, éste, aquí, ahora, constituyen triviales contraejemplos. Cuando pienso «Yo llego tarde al trabajo» puedo hallarme en el mismo estado mental en el que se halla Enrique cuando lo piensa (imagine el lector, si gusta, que Enrique y yo somos gemelos idénticos), y, con todo, la ocurrencia de la palabra «yo» que se da en mi pensamiento se refiere a mí, y la que se da en el pensamiento de Enrique se refiere a Enrique. Cuando un martes pienso «Llego tarde al trabajo» puedo hallarme en el mismo estado mental que cuando lo pienso un miércoles, pero el tiempo al que se refiere mi verbo «Lego», conjugado temporalmente, es diferente en cada uno de los caos...

Supongamos... que en la Tierra Gemela hay castellano-parlantes (por una especie de accidente milagroso, han evolucionado precisamente como nosotros y hablan un lenguaje idéntico al castellano que hace un par de siglos... Supondré que estas gentes no poseen todavía conocimientos de química daltónica o postdaltónica. De modo que, en particular, no disponen de nociones tales como «H<sub>2</sub>O». Entonces, tal como se usa en la Tierra Gemela, la palabra «agua» no se refiere a agua, sino a ese otro líquido (XYZ, digamos). Pese a todo, no hay diferencias relevantes entre el estado mental de los hablantes de la Tierra Gemela y los habitantes de la Tierra (en 1.750) que puedan dar cuenta de esta diferencia referencial. La referencia es diferente porque la *sustancia* es diferente. Por sí mismo, aislado de la situación completa, **el estado mental no fija la referencia**.

**PUTNAM, H.** (1988:22-23). *Razón, verdad e historia*. Madrid: Tecnos.

Algunos filósofos han formulado objeciones a este ejemplo. Estos filósofos sugieren que, si tal planeta se descubriese alguna vez, uno diría que «Hay dos tipos de agua» y no que nuestra palabra «agua» no se refiere al líquido de la Tierra Gemela..

Es sencillo modificar el ejemplo con el fin de eludir el argumento. En primer lugar, el líquido de la Tierra Gemela no tiene por qué ser tan similar al agua. Supongamos que es realmente una mezcla de 20 por 100 de alcohol de grano y otro 80 por 100 de agua, pero la composición química de la gente de a Tierra Gemela es tal que no se intoxican, ni siquiera notan la diferencia de sabor entre esa mezcla y el H<sub>2</sub>O. ... y hallarse así exactamente en el mismo estado mental que un hablante típico de la Tierra en 1750...

... Uno podría conceder que el estado mental individual de una persona no fija la referencia de uno de sus términos, pero insistiendo en que lo que hace el estado mental global de todos los miembros de la comunidad lingüística.

Esta objeción se topa con la dificultad de que la gente de la Tierra o de la Tierra Gemela podría no haber desarrollado tanta química el 1750. Si con anterioridad al desarrollo de la química, el término tenía en la tierra el mismo significado y la misma referencia que hoy posee (en su uso ordinario) y si, con anterioridad al desarrollo de un conocimiento correspondiente, el término tenía en la Tierra Gemela el mismo significado y la misma referencia que hoy tiene, entonces podemos retroceder hasta ese tiempo anterior en el que los estados mentales colectivos coincidían en todos los aquellos aspectos relevantes para fijar la extensión de «agua», argumentando que pese a esto, la extensión era diferente entonces (como lo es ahora) y, de ahí, que **el estado mental colectivo tampoco fija la extensión**. ¿Hemos de decir que la referencia cambió cuando se perfeccionó la química? ¿Diremos que el término fue usado para referirse a ambos tipos de agua (a pesar de que para nosotros su sabor sea diferente) y que sólo se refiere a tipos distintos después de tal progreso?

**PUTNAM, H.** (1988:23-24). *Razón, verdad e historia*. Madrid: Tecnos.

Si decimos que cuando perfeccionaron (o perfeccionamos) la química ) hasta el punto de poder destilar líquidos, de establecer que el agua más el alcohol componen una mezcla, etc.) cambió la referencia de sus términos (o de los nuestros) entonces tendremos que admitir que todos los descubrimientos científicos cambian la referencia de nuestros términos, Pero, según este punto de vista, nosotros no descubrimos que el agua (en el sentido precientífico) era H<sub>2</sub>O; más bien lo *estipulamos*. Esta opinión me parece un caso error. Lo que nosotros queríamos decir con la palabra «agua» abarca desde un principio todo aquello que tuviese la misma naturaleza que la sustancia local *identificada* por ese término; y nosotros descubrimos que el agua, en ese sentido, era H<sub>2</sub>O; lo que la gente de la Tierra Gemela quería decir con la palabra «agua» abarcaba desde un principio la sustancia de su entorno identificada por ese término, y sus expertos descubrieron que el «agua», en ese sentido, era una mezcla de dos líquidos. Si estamos de acuerdo en que la palabra«agua» **no cambia de significado** (en uno u otro lenguaje) cuando los expertos llevan a cabo descubrimientos tales como «el agua es H<sub>2</sub>O» o « el agua es una mezcla de dos líquidos», **o no cambia su significado y su referencia ordinarios** (como resultado de tales descubrimientos pueden parecer más usos técnicos, por supuesto), y que la palabra «agua», en su significado y referencia ordinarios y terrestres, no incluye mezclas de alcohol y de agua, entonces debemos concluir que lo que da cuenta de la diferencia de significado del término «agua» en la Tierra y en la Tierra Gemela no es el conocimiento experto. ... **La palabra «agua» se refería a una sustancia diferente aun cuando el estado mental colectivo de las dos comunidades fuese el mismo. Lo que ocurre en el interior de las cabezas de las personas no fija la referencia de sus términos. Con una frase debida a Mill, «la sustancia misma» completa la tarea de fijar la extensión del término.**

**Una vez vemos que el estado mental** (en el sentido individual o colectivo) **no fija la referencia, deja de sorprendernos el que los cerebros en una cubeta no logren referirse a objetos externos** (aunque se hallen en nuestros mismos estados mentales), y que, por tanto, no puedan decir o pensar que son cerebros en una cubeta.

**PUTNAM, H. (1988:24-25).** *Razón, verdad e historia*. Madrid: Tecnos.  
**INTENCIONES, EXTENSIONES Y «MUNDOS NOCIONALES»**

Puestos a examinar el problema de cómo se fija la referencia de nuestros términos, **dado que no lo fijan nuestros estados mentales**, será conveniente disponer de algunos términos técnicos. **En lógica, el conjunto de cosas de las que un término es verdadero se llama *extensión* del término**. Así pues, la extensión del término «gato» es el conjunto de los gatos. Si un término tiene más de un sentido, entonces simulamos que el término lleve índices invisibles (de modo que tenemos realmente dos palabras y no una), por ejemplo «conejo» —extensión: conjunto de conejos; y «conejo<sub>2</sub>» —extensión: conjunto de los cobardes...

Una palabra como «yo», que se refiere a personas distintas en distintas ocasiones, no tendrá una extensión, sino una función extensión: esto es, una función que determina una extensión en cada contexto de uso. En el caso de la palabra «yo», la función-extensión es bastante sencilla; es simplemente la función  $f(x)$ , cuyo valor para cualquier hablante  $x$  es el conjunto que consiste solamente en  $x$ . En semántica llamamos índice al argumento « $x$ », que rangea sobre todos aquellos parámetros *relevantes* usados en la descripción del contexto (en este caso, el hablante). Los índices son importantes para las cosas a las que nos referimos ostensivamente, y, en el tratamiento semántico completo, también para las características del contexto...

**El conjunto de cosas que componen la extensión de «gato» es distinto en las diferentes situaciones posibles o «mundos posibles».** En un mundo posible  $M$  en que no hay gatos, la extensión de «gato» es el conjunto vacío...



**PUTNAM, H.** (1988:25). *Razón, verdad e historia*. Madrid: Tecnos.

Podemos indicar la forma en que la extensión de un término varía con el mundo posible  $M$  del mismo modo que indicamos que la extensión de la palabra «yo» varía con el hablante: usando una función. Supondremos un conjunto de objetos abstractos, llamados «mundos posibles», que representan los diversos estados de hecho o posibles descripciones del mundo, y asociamos con el término «gato» una función  $f(M)$  cuyo valor en cada mundo posible  $M$  es el conjunto de objetos posibles que son gatos en el mundo  $M$ . Siguiendo a Carnap y Montague, llamaré a esta función *intensión* de la palabra «gato». Similarmente, la intención del predicado diádico «toca», es la función  $f(M)$ , cuyo valor en cada mundo posible  $M$  **es el conjunto de pares ordenados de objetos posibles que el mundo  $M$  se tocan entre sí; la intención del predicado triádico «x está entre y y z» es la función cuyo valor en cualquier mundo posible es el conjunto de triplos ordenados  $(x,y,z)$  tales que  $x$  está entre  $y$  y  $z$  y así sucesivamente.** La intención de una palabra como «yo», cuya extensión depende del contexto, en cada mundo posible, será una función más complicada que tendrá como argumentos los mundos posibles y los índices que representan el contexto.

**La intención especifica el mundo en que la extensión depende del mundo posible. Representa pues lo que nos interesa, la extensión asociada a un término, de un modo muy completo, ya que nos dice qué extensión podría tener en cualquier mundo posible.**

**PUTNAM, H.** (1988:25-26). *Razón, verdad e historia*. Madrid: Tecnos.

La razón por la que la «intensión» (en este sentido) no puede identificarse con el significado, es que cualesquiera dos términos lógicamente equivalentes coinciden extensionalmente en cada mundo posible, teniendo por tanto la misma intensión; pero una teoría que sea incapaz de distinguir entre términos que tienen el mismo significado y términos que son sólo lógica y matemáticamente equivalentes no es una teoría del significado adecuada. «Cubo» y «poliedro regular con seis caras cuadradas» son predicados lógicamente equivalentes. Así pues, la intensión de estos términos es la misma, a saber, la función cuyo valor en cualquier mundo posible es el conjunto de cubos es ese mundo. Pero hay cierta diferencia de significado que podría perderse si identificamos simplemente el significado con la función.

Permítaseme poner énfasis en que, en esta teoría, **los mundos posibles, los conjuntos y funciones han de concebirse como entidades abstractas extramentales y no confundirse con representaciones o descripciones de estas entidades.**

Frege pensó que el significado (Sinn) de una expresión era una entidad extramental o concepto, que podía ser «aprehendido» por la mente, no se sabe cómo. **Esta teoría no nos sirve para el caso de las intensiones**, entendidas en *nuestro sentido*. En primer lugar, como ya apunté, **hay diferencias en el significado que no pueden ser captadas por la intensión**; por ende, **comprender un término no puede consistir sólo en asociarlo con una intensión. Y lo que es más importante, si suponemos que hay un «sexto sentido» que nos permita percibir directamente entidades extramentales (quizá «intuirlas»), entonces la «aprehensión» de una intensión, o cualquier entidad extramental, debe de estar mediada de algún modo por representaciones.**

**PUTNAM, H.** (1988:26-27). *Razón, verdad e historia*. Madrid: Tecnos.

Pero el problema que estamos investigando es precisamente cómo las representaciones pueden permitir que nos refiramos a lo que se halla fuera de la mente. **Dar por sentada la facultad de «aprehender» un x que es exterior a la mente sería cometer una petición de principio.**

Al decir que alguien «cree que hay un vaso de agua en la mesa», lo que hago normalmente es atribuirle la capacidad de referirse al agua... no versa sobre lo que ocurre en la cabeza de Juan, sino que es, en parte, un enunciado sobre el entorno de Juan y sobre su relación con el entorno...

Husserl introdujo un recurso que es útil **cuando deseamos hablar de lo que ocurre en nuestra cabeza sin tener que admitir ninguna presuposición con respecto a la existencia o la naturaleza de las cosas reales a las que se refieren los pensamientos: el recurso de la puesta en paréntesis**. Si «ponemos entre paréntesis» la creencia que adscribimos a Juan cuando decimos «Juan cree que hay un vaso de agua en la mesa», entonces simplemente le adscribimos el estado mental de una persona real o posible que cree que hay un vaso de agua en la mesa (en el pleno sentido ordinario y «sin poner entre paréntesis»). De modo que si Juan no puede notar la diferencia de sabor entre el agua y el agua-con-alcohol-de-grano de la Tierra Gemela, puede hallarse en el mismo estado mental que un hablante real o posible de la Tierra cuando dice «Hay un vaso de agua en la mesa», a pesar de que Juan se refiera con la palabra «agua» a algo que constituirá un cocktail algo cargado. **Diremos que tiene creencia «puesta entre paréntesis» de que [hay un vaso de agua en la mesa]**. En efecto, **el recurso de la puesta entre paréntesis resta implicaciones de la locución ordinaria (todas las implicaciones de la locución ordinaria de creencia(todas las implicaciones que se refieren al mundo externo, o a lo que es exterior a la mente del sujeto pensante).**

**PUTNAM, H.** (1988:27). *Razón, verdad e historia*. Madrid: Tecnos.

Daniel Dennett ha utilizado recientemente la locución «**mundo nocional**» de un modo afín a como Husserl utilizó la puesta entre paréntesis. **La totalidad de las creencias puestas entre paréntesis del sujeto pensante constituye la descripción de su mundo nocional**, en el sentido de Dennett. Así pues la gente de la Tierra Gemela aproximadamente tiene el mismo *mundo nocional* ... que tenemos nosotros; lo único que ocurre es que viven un planeta real diferente y (usan «agua» para referirse a una sustancia real distinta); y los cerebros de una cubeta ... podrían haber tenido el mismo mundo nocional que nosotros tenemos, hasta el último detalle; lo único que pasa es que ninguno de sus términos tenía referencia en el mundo exterior. **La teoría tradicional del significado daba por sentado que el mundo nocional del sujeto pensante determinaba las intensiones de sus términos** (y éstas, junto con el hecho de que el mundo posible particular M es el actual, **determinan las extensiones de los términos y los valores veritativos de las oraciones**). Hemos visto que la teoría tradicional del significado es incorrecta; por eso la literatura contiene hoy muchos conceptos diferentes (por ejemplo «intensión» y «mundo nocional»), y no un solo concepto unitario de «significado». **El «significado» se ha hecho pedazos.... Si la intención y la extensión no están directamente determinadas por el mundo nocional, ¿cómo están determinadas entonces?**

**PUTNAM, H.** (1988:27-28). *Razón, verdad e historia*. Madrid: Tecnos.

## **LA CONCEPCIÓN ADMITIDA CON RESPECTO A LA INTERPRETACIÓN**

La opinión más común con respecto al modo en que fijamos las interpretaciones de nuestro lenguaje (si no individual, al menos colectivamente), está asociada con las nociones de *constreñimiento operacional* y *constreñimiento teórico*. **Los constreñimientos operacionales se concibieron originalmente de un modo bastante ingenuo; simplemente estipulamos** (convencionalmente, como fuera así) que cierta oración (digamos «fluye electricidad a través del alambre») **es verdadera si y sólo si se observa cierto resultado en una prueba....** Esta especie de crudo operacionalismo ya no tiene defensores, pues ya se ha apreciado que (1) **los vínculos entre teoría y experiencia son probabilísticos y no pueden ser correctamente formalizados como perfectas correlaciones** (aun cuando fluya electricidad a través del alambre, hay siempre sucesos de baja probabilidad o condiciones de fondo que podrían impedir las oscilaciones de la aguja del voltímetro); y (2) estos vínculos ni siquiera son simples correlaciones semánticas, sino que dependen de una teoría empírica que está sujeta a revisión. Según el operacionalismo ingenuo, cada vez que se descubre un nuevo test para comprobar si una sustancia es realmente «oro», el significado y la referencia de oro sufre un cambio. (De hecho, no diríamos que se descubre un nuevo criterio para reconocer el oro). Desde una perspectiva operacionalista, las teorías son contrastadas con la experiencia oración por oración (los significados operacionales estipulados para las oraciones individuales nos dicen cómo emprender la contrastación de la teoría); Según una descripción más reciente, **las teorías se someten al tribunal de la experiencia como cuerpos integrados, como expuso Quine.**

**PUTNAM, H.** (1988:28). *Razón, verdad e historia*. Madrid: Tecnos.

Sin embargo, es posible relajar la noción de constreñimiento operacional con el fin de superar estas objeciones. Así, uno podría restringir la clase de **interpretaciones (asignaciones de intensiones a los predicados de su lenguaje) admisibles**, de acuerdo con constreñimientos de la forma «una interpretación es admisible si la mayoría de las veces la oración S es verdadera cuando se satisface la condición experimental E» (respectivamente, si la mayoría de las veces la oración S es falsa cuando se satisface E). Tales constreñimientos conforman la idea de que hay relaciones probabilísticas entre la experiencia y la verdad o falsedad de las oraciones del lenguaje. .. En vez de concebirlos como estipulaciones de significado, a la manera del operacionalismo crudo, pueden concebirse como tentativas de restricción de la clase de interpretaciones admisibles...En resumen, puede adoptarse la opinión de que lo que selecciona la interpretación de nuestros términos son **los constreñimientos operacionales que los investigadores racionales impondrían, si observasen y experimentasen razonando tan correctamente como fuera posible: los constreñimientos que adoptarían en el estado de «equilibrio reflexivo»; los constreñimientos que de hecho aceptamos en un momento dado tienen el *status* de estimación racional o aproximación.**

**PUTNAM, H.** (1988:28-29-30). *Razón, verdad e historia*. Madrid: Tecnos.

Además de restringir la clase de interpretaciones admisibles por medio de constreñimientos operacionales (o mediante aproximaciones sucesivas al conjunto ideal peirciano de constreñimientos operacionales), también podemos disponer de constreñimientos que se remitan a propiedades formales de la teoría. Por ejemplo, «Una interpretación admisible es aquella en la que resulta ser cierto que diferentes efectos tienen diferentes causas». Kant sostuvo que tal «constreñimiento teórico» formaba parte de la propia racionalidad: imponemos en el mundo el principio del determinismo, más bien que descubrirlo. Esta formulación del principio es demasiado enfática: el precio de preservar el determinismo podría ser una complicación excesiva de nuestro sistema cognitivo, considerado como un todo.

... «El conjunto de oraciones verdadero bajo una interpretación no debe tener un grado inferior de simplicidad que cualquier otro conjunto que posea las mismas consecuencias observacionales y experimentales» correspondería en lógica inductiva al constreñimiento que conmina a aceptar la hipótesis más simple (o la más «plausible») entre aquellas que son compatibles con las observaciones que se llevan a cabo.

... **Los motivos que hacen atractiva la idea de que las interpretaciones de nuestro lenguaje (en el sentido de asignaciones de intensiones a sus términos) se fijan por medio de constreñimientos operacionales teóricos, son obvios:** la mente puede discernir si se están teniendo o no cierto tipo de experiencia ... De forma que si una teoría implica o contiene una oración asociada con una experiencia E por medio de algún tipo de constreñimiento operacional, probabilístico, o de cualquier otra índole, entonces el sujeto pensante puede saber si la teoría es operativa, o por el contrario, hay una dificultad de ajuste —al menos en ese caso— observado si tiene o no experiencia E.

**PUTNAM, H.** (1988:28-30). *Razón, verdad e historia*. Madrid: Tecnos.

Y dado que los constreñimientos utilizados para contrastar la teoría fijan también las extensiones de los términos, lo que el sujeto pensante estime sobre la «operatividad» de la teoría constituye también una estimación sobre la verdad de ésta. Y ya que el hablante, al conocer esos constreñimientos, conoce también las intensiones de los términos, se sigue que la aprehensión de una semántica correcta le informaría cómo ha de ser el mundo para que la teoría —cualquier teoría del mundo propuesta— sea verdadera...

En resumen, si la concepción admitida fuese correcta, tendríamos una elegante descripción de cómo se **fijan las intensiones y las extensiones** (en principio, pues es difícil atar todos los cabos, dado el estado actual de nuestro conocimiento metodológico). Pero, por desgracia, la concepción admitida no resuelve el problema.



**PUTNAM, H.** (1988:**30-31**). *Razón, verdad e historia*. Madrid: Tecnos.

## **POR QUÉ FRACASA LA CONCEPCIÓN ADMITIDA**

La dificultad de la concepción admitida es que intenta fijar las intensiones y las extensiones de los términos individuales, fijando las condiciones de verdad de las oraciones completas. Como ya vimos, la idea es que los constreñimientos operacionales y teóricos (aquellos que los investigadores racionales aceptarían en algún tipo de límite ideal de investigación) determinan qué oraciones son verdaderas en el lenguaje. Sin embargo, aunque esto sea correcto, tales constreñimientos no pueden determinar la referencia de nuestros términos...

Tal fracaso ha sido mostrado por QUINE... Resumiendo, no sólo es que fracase la concepción admitida: ningún criterio que únicamente fije los valores de verdad de oraciones completas puede fijar la referencia, incluso si especifica los valores de verdad de las oraciones en cada mundo posible.

**PUTNAM, H.** (1988:36). *Razón, verdad e historia*. Madrid: Tecnos.

## **«SUPERVIVENCIA» Y EVOLUCIÓN**

La sugerencia que hoy en día está de moda es que el mismo proceso evolutivo ha producido, de un modo u otro, una correspondencia entre nuestras palabras, las representaciones mentales y las cosas externas; se dice que no habríamos sobrevivido si no se hubiese dado tal correspondencia, y que ésta constituye la relación de referencia, por lo menos en un nivel rudimentario.

Pero, ¿qué tienen que ver «correspondencia» y «referencia» con la *supervivencia*? Y, de paso, ¿qué tiene que ver la verdad con la supervivencia?

...Algunos filósofos creen que no sobreviviríamos si (una cantidad suficiente de ) nuestras creencias no fuesen *verdaderas*. Otros afirman que ni siquiera nuestras creencias científicas mejor establecidas son verdaderas, o al menos que no tenemos razón alguna para pensar que lo son. **Thomas Kuhn ha sugerido que nuestras creencias sólo se «refieren» a objetos dentro de esas creencias** ... el éxito de la ciencia se explica por ensayo-error y no por una correspondencia entre sus objetos y cosas reales, apunta Kuhn. En su nuevo libro, Bas van Fraassen sostiene que una teoría exitosa no necesita ser verdadera, sino sólo «observacionalmente adecuada», esto es, que prediga correctamente la observación. También explica el éxito (o adecuación observacional) de la ciencia como producto del ensayo-y-error.

**PUTNAM, H.** (1988:36-37). *Razón, verdad e historia*. Madrid: Tecnos.

Si estos filósofos están en lo cierto, la idea de emplear la evolución para justificar la creencia en una relación *objetiva* de referencia es gratuita. Según la perspectiva instrumentalista, la evolución únicamente establece una correspondencia entre algunos términos (los términos observacionales) y «posibilidades» permanentes de sensación. Tal correspondencia no puede ser la *referencia*, a menos que deseemos abandonar la idea de que las cosas externas (las observables) son algo más que menos constructos a partir de sensaciones.

No obstante, **creo que son los otros filósofos quienes están en lo cierto (aquellos que dicen que no sobrevivimos si una cantidad suficiente de nuestras creencias no fuesen verdaderas).**

**Y lo creo por una razón: el ensayo-y-error no explica por qué nuestras teorías son «observacionalmente adecuadas»;** tal explicación solo puede darse en relación con las características de la *interacción* hombre-medio, siendo éstas las que explican el éxito del ensayo-y-error ¡El ensayo-y-error no tiene éxito en todas las empresas, después de todo!). **Postular que la interacción ocasiona en nuestras mentes teorías falsas cuyas consecuencias son precisamente predicciones exitosas es postular una serie de coincidencias totalmente inexplicable.**

**Algunas de nuestras creencias están íntimamente ligadas con la acción.** Si creo en la oración «Si pulso este botón conseguiré algo de gran valor para mí» (suponiendo que entiendo esta oración de un modo normal, o al menos le asocio la creencia «puesta entre paréntesis» o «noción» normal), entonces alargaré la mano y pulsaré el botón. Llamo **creencias directivas a las creencias de la forma: «Si hago x, conseguiré...» donde el espacio en blanco representa una meta del agente.. De modo que la verdad de (una cantidad suficiente de) nuestras creencias *directivas* es necesaria para la supervivencia.**

**PUTNAM, H.** (1988:37-38). *Razón, verdad e historia*. Madrid: Tecnos.

Ahora bien, nuestras creencias *directivas* se derivan a su vez de otras muchas creencias: creencias sobre las características y sobre las facultades causales de las cosas externas, y creencias sobre nuestras propias características y facultades. Si la mayor parte de estas creencias fuesen falsas, ¿no sería una mera coincidencia que, pese a ello, condujesen a predicciones de la experiencia y a creencias directivas verdaderas? Así pues, ya que (una cantidad suficiente de) nuestras creencias directivas son verdaderas, y puesto que la mejor explicación de este hecho es que muchas de nuestras restantes creencias (las que constituyen nuestra «teoría del mundo cotidiano») son verdaderas, al menos aproximadamente, tenemos buenas razones para creer que nuestra teoría del mundo cotidiano es verdadera, al menos aproximadamente, y que si no fuera así no hubiésemos sobrevivido.

... Ahora bien, si una cantidad suficiente de nuestras creencias directivas son verdaderas bajo la interpretación no-estándar J, entonces desde luego, obtendremos éxito y sobreviviremos ... En resumen, la J-verdad de (una cantidad suficiente de ) creencias directivas es tan óptima para el «éxito evolutivo» como la J-verdad, ya que las condiciones de verdad de cada oración (no sólo de las creencias directivas) son las mismas bajo J y bajo J. **Mis creencias directivas no sólo están asociadas con la misma experiencia subjetiva bajo la interpretación J y bajo la interpretación J; tienen además las mismas condiciones de verdad.** Desde el punto de vista (o no-punto de vista) de la «evolución», **todo lo que se necesita es que una cantidad suficiente de mis creencias sean verdaderas bajo cualquier interpretación que conecte tales creencias con acciones relevantes. La evolución puede provocar en mí cierta tendencia a tener creencias verdaderas (de ciertos tipos); pero** esto únicamente quiere decir que la evolución influye lingüística y conceptualmente sobre la supervivencia mediante su tendencia a generar en nosotros sistemas de representación cuyas oraciones, o análogos de oración, poseen ciertas condiciones de verdad (y ciertas condiciones de acción o «reglas de salida del lenguaje»). Pero mostramos que las condiciones de verdad para las oraciones completas no determinan la referencia de sus partes... **Se sigue que es sencillamente un error pensar que la evolución determina una única correspondencia ... entre las expresiones referenciales y los objetos externos.**

**PUTNAM, H.** (1988:**38**). *Razón, verdad e historia*. Madrid: Tecnos.

## **INTENCIONES: PURAS E IMPURAS**

Hemos visto que la naturaleza no selecciona ninguna correspondencia entre nuestros términos y las cosas externas. La naturaleza nos conmina a tratar las palabras y los signos mentales de tal modo que una cantidad suficiente de nuestras creencias directivas resulten verdaderas, a fin de que una cantidad suficiente de nuestras acciones contribuyan a nuestra «aptitud genética completa»; pero esto deja la referencia en gran parte indeterminada. W.V.Quine ha insistido en que esto es, de hecho, lo que es **la referencia es: indeterminada**. El autor de *La Relatividad Ontológica* piensa que es una quimera considerar que los términos de nuestro lenguaje tengan contrapartidas bien definidas. Con sus palabras:

Consideremos de nuevo nuestra notación canónica, con un léxico de predicados interpretados y algún rasgo de valores fijado para las variables de cuantificación. Las oraciones verdaderas de este lenguaje continúan siendo verdaderas bajo incontables reinterpretaciones de los predicados e incontables revisiones del rango de valores de las variables. En realidad, podemos hacer que sirva cualquier rango del mismo tamaño por medio de una adecuada reinterpretación. Si el rango de valores es infinito, podemos hacer que sirva cualquier rango infinito; éste es el Teorema de Skolem-Lowenheim. Las oraciones verdaderas continúan siendo verdaderas bajo todos los cambios.

PUTNAM, H. (1988:38-39). *Razón, verdad e historia*. Madrid: Tecnos.

La alternativa aquí sugerida, y que examinaré ... es abandonar la idea que hasta aquí ha servido como premisa en toda discusión: **que nuestras palabras están en una especie de relación uno-a-uno con las cosas (independientes del discurso) y con conjunto de cosas. Sin embargo puede parecer que hay una solución mucho más simple: ¿por qué no decir que son precisamente nuestras intenciones, implícitas o explícitas las que fijan la referencia de nuestros términos?**

Nada más comencé esta discusión en el capítulo rechacé esta opción sobre la base de que no constituía una respuesta informativa, puesto que **el tener intenciones (de tipo relevante) presupone ya la capacidad de referir.**

**El problema es que las nociones de «intención» y «estado mental» poseen cierta ambigüedad.** Llamemos **puro** a un estado mental si su presencia o ausencia depende sólo de lo que acaece dentro del hablante. El que yo tenga o no dolor sólo depende de lo que acaece «dentro de mí», pero que sepa o no que la nieve es blanca depende no sólo de que algo acaezca o no «dentro» de mí (creer o confiar en que la nieve es blanca), sino que también depende de que la nieve sea o no blanca, y, de este modo, de algo que acaece «fuera» de mí cuerpo y de mi mente. Así pues, el dolor es un estado mental puro, pero el conocimiento es un estado mental **impuro**. Hay un componente mental (puro) en el conocimiento, pero también hay un componente que de ningún modo es mental: aquél que corresponde a la condición de que lo que un hombre cree no constituye conocimiento a menos que la creencia sea verdadera. No me hallo en el «estado» de conocer que la nieve es blanca si no estoy en un estado mental puro conveniente. Pero **estar en un estado mental puro conveniente nunca basta para saber que la nieve es blanca; el mundo tiene que cooperar también.**

¿Y la creencia? Hemos definido la **creencia puesta entre paréntesis («mundo nocional»)** de forma que tener una creencia puesta entre paréntesis de que [hay agua en la mesa], o tener un mundo nocional que incluya el haber agua sobre la mesa, es un estado mental puro. Pero ... el creer que hay agua en la mesa (sin poner nada «entre paréntesis») presupone que la palabra «agua» se refiere realmente a agua, y esto depende de la naturaleza real de ciertos «paradigmas», de las relaciones causales directas que se mantengan con estos paradigmas, etc. Cuando tengo la creencia de que hay agua en la mesa, mi doble en la Tierra Gemela tiene la misma creencia puesta entre paréntesis, pero no la misma creencia ya que su palabra «agua» se refiere al agua-con-alcohol-de-grano y no al agua. **En resumen, creer que hay agua en la mesa es un estado mental impuro** (los cerebros en una cubeta no podrían estar en tal estado, aunque si en el correspondiente estado «puesto entre paréntesis»).

**A la intención le ocurre lo mismo que a la creencia.** Los estados mentales e intencionales puros —por ejemplo, **el acto intencional de que el término «agua» se refiere al agua en el mundo nocional de uno— no fija la referencia en el mundo real.** Los estados mentales intencionales impuros —el acto intencional de que el término «agua» se refiera al agua real— presuponen la capacidad de referirse al agua (real).

**PUTNAM, H.** (1988:39-40). *Razón, verdad e historia*. Madrid: Tecnos.

Algunos filósofos han sugerido que la creencia puede definirse en términos del estado que llamé «creencia puesta entre paréntesis»:

Juan cree que la nieve es blanca = Juan cree que [la nieve es blanca].

(Esto es, la nieve es blanca en el mundo nocional de Juan).

Y en el pensamiento de Juan, las palabras «nieve» y «blanca» (o cualesquiera palabras que use para expresar su creencia) se refieren a la nieve y a la propiedad de ser blanca, respectivamente.

Sin admitir esta descripción como un análisis correcto y completo de lo que es creer que la nieve es blanca, podemos aceptar que pone de relieve un punto que sin duda es correcto: **creer presupone la capacidad de referir**. Y exactamente del mismo modo, **los actos intencionales ya presuponen la capacidad de referir. Las intenciones no son acontecimientos mentales que causan la referencia de las palabras: las intenciones (en el sentido ordinario e impuro) contienen la referencia como un componente integral**. Explicar la referencia en términos de intención (impura) sería circular. Y el problema de cómo los estados mentales *puros intencionales*, de creencia, etc. , pueden (en el marco causal apropiado) constituir o causar la referencia es precisamente lo que hemos hallado demasiado enigmático.



**PUTNAM, H.** (1988:**40**). *Razón, verdad e historia*. Madrid: Tecnos.

## **EL ORIGEN DEL ENIGMA**

A primera vista parece sumamente obvio que nuestras palabras y representaciones mentales refieren. Cuando pienso o digo «El gato acaba de salir», usualmente pienso algo acerca de nuestro gato Mitty: la palabra gato, en esa oración, pensada o dicha, se refiere a un conjunto de entidades del que es miembro Mitty. A pesar de todo, acabamos de ver que la naturaleza de esta relación de «acerca de qué» o referencia es enigmática.

La distinción entre el mundo real y el mundo nocional (y la correspondiente distinción entre creencias y puestas entre paréntesis) explica parte del enigma. **La razón por la que resulta sorprendente y problemático descubrir que hay «interpretaciones admisibles» no intencionales de nuestro lenguaje (entiendo por interpretación admisible una interpretación que satisface los constreñimientos operacionales y teóricos apropiados) es, en parte, que en el «mundo nocional» del hablante no surge tal «indeterminación».** En mi mundo nocional los gatos y los gatos\* son completamente distintos (de hecho, en mi mundo nocional los gatos\* son cerezas). «Hay un gato en una estera» y «Hay un gato\* en una estera\*» pueden ser lógicamente equivalentes, pero contienen términos con referentes nocionales totalmente distintos; por eso parece realmente extraño que pueda haber confusión entre los referentes del mundo real de una creencia y los de otra.

**PUTNAM, H.** (1988:**40-41**). *Razón, verdad e historia*. Madrid: Tecnos.

## EL ORIGEN DEL ENIGMA

Pero si el número de gatos resulta ser igual que el número de cerezas, entonces se sigue que los teoremas de la Teoría de Modelos (como Quine subraya en la cita anterior) que existe una reinterpretación de todo el lenguaje que no altera los valores veritativos de todas las oraciones aunque se permuten las extensiones de «gato» y «cereza» : gracias a las técnicas ya mencionadas, tales interpretaciones pueden concebirse de manera que preserven los constreñimientos operacionales y teóricos (y mediante las técnicas ejemplificadas en el caso «gato/gato\*», pueden ampliarse hasta que nos proporcionen «intensiones», o funciones que determinen una extensión en cada mundo posible, y no sólo en el mundo actual). Ello no contradice los enunciados que establecimos sobre nuestro mundo nocional, o sistema subjetivo de creencias, por la siguiente razón: el hecho de que en nuestro sistema de creencias o «mundo nocional» ningún gato sea una cereza significa que en cada interpretación de tal sistema (en cada asignación de referentes del mundo externo a los términos, a las imágenes y a las demás representaciones que empleamos al pensar), los referentes de «gato» y los referentes de «cereza» deben ser conjuntos distintos. Pero la disyuntividad de tales conjuntos es comparable con el hecho (subrayable) de que lo que es un conjunto de los «gatos» en una interpretación admisible puede ser el de las «cerezas» en una interpretación diferente (pero igualmente admisible).

**PUTNAM, H.** (1988:41). *Razón, verdad e historia*. Madrid: Tecnos.

**Del hecho de que los gatos nocionales sean distintos de las cerezas nocionales sólo se sigue que los gatos reales son completamente distintos de las cerezas reales si el número de interpretaciones es exactamente uno. Si hay más de una interpretación admisible** de todo lenguaje ... entonces dos términos que se refieren a conjuntos distintos en cada interpretación admisible, **pueden tener los mismos referentes potenciales cuando se tiene en cuenta la totalidad de interpretaciones admisibles**. Del hecho de que los gatos nocionales sean diferentes a las cerezas nocionales no se sigue que haya conjuntos disjuntos de gatos-en-si-mismos y cerezas-en-si-mismas; y y si esto resulta tan penoso es porque **los constreñimientos operacionales más los teóricos son el medio natural para permitir que el contexto empírico determine la interpretación (o las interpretaciones) admisibles del sistema representacional de un sujeto**. Tales constreñimientos pueden determinar, hasta cierto punto, las oraciones que son verdaderas y las que son falsas en nuestro lenguaje; y ésta es la escasa actividad que queda entre las condiciones de verdad y la referencia. Quine desearía acabar con esta relación mórbida y reconocer simplemente que la referencia queda ideterminada.

**PUTNAM, H.** (1988:**41-42**). *Razón, verdad e historia*. Madrid: Tecnos.

... Harty Field, un joven filósofo, ha sugerido una alternativa distinta. En opinión de Field, la referencia es una «relación fisicalista», es decir, una relación causal compleja entre palabras o representaciones mentales y objetos o conjuntos de objetos. Según Field, corresponde a la ciencia empírica descubrir lo que sea tal relación fisicalista. No obstante, esta sugerencia también tropieza con problemas. Supongamos que, como Field sostiene, hay una posible definición fisicalista naturalista de la referencia.

Supongamos que:

(1) x se refiere a y si y sólo si x mantiene **R** con y es verdadera, donde **R** es una relación definible con el vocabulario de la ciencia natural, sin usar nociones semánticas (es decir, sin usar «referente» o cualesquiera otras palabras que harían la definición inmediatamente circular). Si (1) es verdadera y empíricamente verificable, entonces (1) es una oración verdadera hasta en la teoría según la cual la referencia se fija en tanto que (y sólo en tanto que) es determinada por los constreñimientos operacionales más teóricos. Así pues (1) es una oración que formaría parte de nuestra teoría del mundo en el «límite ideal» o «equilibrio reflexivo».

**PUTNAM, H.** (1988:42). *Razón, verdad e historia*. Madrid: Tecnos.

Sin embargo, si la referencia sólo está determinada por los constreñimientos operacionales y teóricos, entonces la referencia de «x mantiene R con y» está ella misma indeterminada, y por ende, saber que (1) es verdadera será de poca ayuda. Cada modelo admisible de nuestro lenguaje-objeto tendrá su correspondiente modelo en nuestro meta-lenguaje, en el cual (1) se cumple. La interpretación de «x mantiene R con y» fijará la interpretación de «x se refiere a y». **Pero ésta será sólo una relación en cada modelo admisible; de ninguna manera servirá para acotar el número de interpretaciones admisibles.**

Field no pretende afirmar esto, desde luego. Field está afirmando que (a) hay una relación única entre palabras y cosas o conjuntos de cosas; y (b) ésta es la relación que debemos usar como relación de referencia al asignar un valor veritativo a (1). Pero, como acabamos de ver, no expresamos necesariamente esto sólo con decir (1); y cómo podríamos aprender a expresar que lo que Field quiere decir es un misterio.

Dejando de lado este último enigma, consideremos la opción de que (1), si la entendemos como Field quiere que lo hagamos (como si describiese la relación única y determinada entre palabras y sus referentes), es verdadera. Si, entendida de este modo (1) es verdadera, ¿qué la hace verdadera? Dado que hay muchas «correspondencias» entre palabras y cosas, e incluso muchas correspondencias que satisfacen nuestros constreñimientos, ¿qué *selecciona* a una particular correspondencia R? No es una correlación empírica de (1); porque esto es un asunto de nuestros constreñimientos operacionales y teóricos. **No son, como hemos visto, nuestras intenciones (más bien R toma parte en la determinación de lo que nuestras intenciones significan).** Parece que el hecho que R sea la referencia debe ser un hecho *metafísicamente inexplicable*, un tipo de verdad metafísica primitiva, irracional.

**PUTNAM, H.** (1988:42-43). *Razón, verdad e historia*. Madrid: Tecnos.

La indicación de Kripke, que está íntimamente relacionada con las anteriormente realizadas sobre la referencia de los términos de géneros naturales (por ejemplo, términos para especies animales, vegetales y minerales), consistía en que, dado que como cuestión de hecho

(2) El agua es H<sub>2</sub>O

(es decir, dado que (2) es verdadera en el mundo real) y dado que (señala Kripke) los hablantes efectúan el **acto intencional** de que el término «agua» se refiera sólo a aquellas cosas que tienen la misma conducta y la misma composición última que los diversos estándares del agua real (esto es, los hablantes tienen tales **intenciones incluso cuando hablan sobre casos hipotéticos o «mundos posibles»**), se sigue que (2) también debe ser verdadera en todo mundo posible; porque describir un hipotético líquido que no es H<sub>2</sub>O, pero que tiene algunas características similares a las del agua, es sólo describir un hipotético líquido que se asemeja al agua, y no describir un mundo posible en el que el agua no es H<sub>2</sub>O; pero esta «necesidad metafísica» se explica por la química y por los hechos mundanos **sobre las intenciones referenciales de los hablantes**.

Incluso si hay una relación fisicalista determinada R ... que sea precisamente la referencia.. este mismo hecho no puede ser consecuencia de nuestras intenciones referenciales; más bien, tal y como hemos apuntado repetidas veces, toma parte en la determinación de lo que **nuestras mismas intenciones referenciales significan**. La opinión de Kripke según la cual «El agua es H<sub>2</sub>O» es verdadera en todos los mundos posibles, podría ser acertada incluso si la referencia en el mundo real se fijase sólo por constreñimientos operacionales y teóricos su punto de vista presupone la noción de referencia, no nos dice si la referencia está determinada o qué es la referencia.

**PUTNAM, H.** (1988:**43**). *Razón, verdad e historia*. Madrid: Tecnos.

En mi opinión, creer que alguna correspondencia es precisa e intrínsecamente la referencia (no como resultado de nuestros constreñimientos operacionales y teóricos, o de nuestras **intenciones**, **sino como un hecho metafísico último**) **viene a ser lo mismo que mantener una teoría mágica de la referencia**. Desde tal perspectiva, la propia referencia se convierte en lo que Locke llamó «forma substancial» (una entidad que pertenece intrínsecamente a cierto nombre). .. Los problemas epistemológicos que acompañan a tal perspectiva metafísica parecen insuperables. Porque **suponiendo un mundo de entidades independientes de la mente y del discurso (y esto es lo que presupone el punto de vista que estamos discutiendo)** **hay como hemos visto, muchas «correspondencias» diferentes que representan relaciones de referencia posible o candidatas** (infinitamente muchas, de hecho, si hay una cantidad infinita de cosas en el universo).

**PUTNAM, H.** (1988:**43**). *Razón, verdad e historia*. Madrid: Tecnos.

Ni siquiera la exigencia de que (1) sea verdadera, bajo la noción de verdad que corresponda a la relación «real» de referencia metafísicamente seleccionada, logra excluir ninguna de estas candidatas, si (1) es empíricamente aceptable (aceptable según nuestros constreñimientos operacionales y teóricos). Más entonces hay una cantidad infinita de posibles «verdades metafísicamente irracionales» diferentes, de la forma «R es la relación de referencia real (metafísicamente singularizada)». **Si el que sostiene tal punto de vista concede que podría no ser del todo correcto, y que la referencia puede estar metafísicamente seleccionada R puede permitir una determinada pluralidad de interpretaciones admisibles**), entonces hasta puede ser concebible que el punto de vista de los constreñimientos-operacionales-y-teóricos sea, metafísicamente correcto, después de todo. Pues, ¿por qué no podría ser un hecho metafísicamente explicable el que la referencia sea la relación: x se refiere a y al menos en un modelo admisible M?



**PUTNAM, H.** (1988:45). *Razón, verdad e historia*. Madrid: Tecnos.

### 3. **DOS PERSPECTIVAS FILOSÓFICAS**

Los problemas que hemos estado discutiendo dan origen de por sí a dos puntos de vista filosóficos...

Una de ellas es la del **realismo metafísico**. Según esta perspectiva, **el mundo consta de alguna totalidad fija de objetos independientes de la mente**. Hay exactamente una descripción verdadera y completa de «cómo es el mundo». La verdad supone una especie de relación de correspondencia entre palabras o signos mentales y cosas o conjuntos de cosas externas. A esta perspectiva la llamaré *externalista*, ya que **su punto de vista predilecto es el del Ojo de Dios**.

La perspectiva que voy a defender carece de nombre que no sea ambiguo. Es un logro tardío en la historia de la filosofía... La denominaré perspectiva **internalista**, ya que lo característico de tal concepción es sostener que sólo tiene sentido formular la pregunta **¿de qué objetos consta el mundo?** Desde dentro de una teoría o descripción. Muchos filósofos internalistas... sostienen además que hay más de una teoría o descripción del mundo verdadera. Desde la perspectiva internalista, **la «verdad» es una especie de aceptabilidad racional (idealizada) —una especie de coherencia ideal de nuestras creencias entre sí y con nuestras experiencias, considerándolas como experiencias representadas en nuestro sistema de creencias— y no una correspondencia con «estados de cosas» independientes de la mente o del discurso**.

**PUTNAM, H.** (1988:**45-46**). *Razón, verdad e historia*. Madrid: Tecnos.

No existe punto de vista como el del Ojo Divino que podamos conocer o imaginar con provecho. Sólo existen diversos puntos de vista de personas reales, que reflejan[*reflecting*] aquellos propósitos e intereses a los que subordinan sus descripciones y teorías. («Teoría de la verdad-coherencia», «no-realismo», «verificacionismo», «pluralismos», «pragmatismo»; pese a que se han aplicado todos estos términos a la perspectiva internalista.. tiene connotaciones inaceptables debido a sus restantes aplicaciones históricas). Los filósofos internalistas rechazan la hipótesis de los «cerebros en una cubeta». La hipótesis de los «cerebros en un mundo-cubeta» es para nosotros únicamente un relato, una mera construcción lingüística: de ningún modo un mundo posible. La idea de que este relato podría ser verdadero en algún universo, en alguna Realidad-Paralela, supone desde el principio el punto de vista del Ojo de Dios, como fácilmente puede verse. En efecto, ¿desde qué punto de vista se cuenta este relato? Evidentemente, no desde el punto de vista de algún observador de otro mundo que interactúe con éste, pues un mundo» incluye por definición todo aquello que interactúa, de una u otra forma, con las cosas que contiene.

**PUTNAM, H.** (1988:46). *Razón, verdad e historia*. Madrid: Tecnos.

Si usted, por ejemplo, fuera el observador que no es un cerebro en una cubeta, espiando a los cerebros en una cubeta, entonces el mundo no sería un mundo en el que todos los seres sintientes fueran cerebros en una cubeta. Sí que la suposición de que podría haber un mundo en el que todos los seres sintientes fueran cerebros en una cubeta presupone desde el principio a visión de la verdad del Ojo Divino —o con más precisión, la visión del No-Ojo—, la verdad como algo totalmente independiente de los observadores.

Como hemos visto, el problema es el siguiente: ahí fuera hay esos objetos. Aquí la mente-cerebro, llevando a cabo su pensamiento-computación. ¿Cómo entran los símbolos del sujeto pensante (o los de su mente-cerebro) en una **correspondencia** única con los objetos de ahí fuera?

Por otra parte, el filósofo externalista no puede rechazar tan fácilmente la hipótesis de que todos somos cerebros en la cubeta. Pues la verdad de una teoría no consiste en su ajuste con el mundo conforme éste se presenta al observador u observadores (la verdad no es «relacional» en este sentido), **sino en su correspondencia con el mundo tal como es en sí mismo**. Y el problema que le planteo al filósofo externalista, es que si él es un cerebro en una cubeta, no puede disponer lógicamente de la misma relación de correspondencia de la que (en su opinión) depende la verdad y la referencia. Por tanto, si somo cerebros en una cubeta, no podemos pensar que lo somos; excepto en el sentido puesto entre paréntesis... **no tiene condiciones referenciales que lo hagan verdadero**. De modo que, después de todo, no es posible que seamos cerebros en una cubeta.

**PUTNAM, H.** (1988:46-47). *Razón, verdad e historia*. Madrid: Tecnos.

La réplica que está hoy de moda entre los externalistas es que, pese a que en realidad ningún signo corresponde *necesariamente* a ningún conjunto de cosas más bien que a otro, las conexiones *contextuales* entre los signos y las cosas externas (en particular, las conexiones causales) harán explicable la naturaleza de la **referencia**. Pero así no se resuelve el problema. Por ejemplo, es probable que la causa predominante de mis creencias sobre los electrones sean diversos manuales. Pero mis preferencias de la palabra «electrón», pese a tener en este sentido una firme conexión con los manuales, *no se refieren a éstos*. **Los objetos que son la causa predominante de mis creencias contenientes de cierto signo pueden no ser los referentes de ese signo.**

El externalista replicará ahora que la palabra «electrón» no está conectada con los manuales mediante una cadena del tipo apropiado. (Pero ¿cómo podemos **tener intenciones que determinen qué cadenas causales son «del tipo apropiado» salvo que ya seamos capaces de referirnos a las cosas?**).

**PUTNAM, H.** (1988:47). *Razón, verdad e historia*. Madrid: Tecnos.

Para un **internalista como yo**, la situación es completamente distinta. Desde una perspectiva internalista, los signos tampoco corresponden intrínsecamente a objetos con independencia de quién y cómo los emplee. Pero un signo empleado de un modo determinado por una determinada comunidad de usuarios puede corresponder a determinados objetos dentro del esquema conceptual de esos usuarios. **Los «objetos» no existen independientemente de los esquemas conceptuales. Desmenuzamos el mundo en objetos cuando introducimos uno u otro esquema descriptivo, y puesto que tanto los objetos como los símbolos son internos al esquema descriptivo, es posible indicar cómo se emparejan.**

En realidad es trivial decir cuál es la **referencia** de alguna palabra dentro del lenguaje al que pertenece mediante el uso de la misma palabra. ¿A qué se refiere «conejo»? ¡Cómo! ¡A conejos, por supuesto! ¿A qué se refiere extraterrestres? A extraterrestres (si los hay).

**PUTNAM, H.** (1988:**47-48**). *Razón, verdad e historia*. Madrid: Tecnos.

Por supuesto, el externalista está de acuerdo en que la extensión de «conejo» es el conjunto de los conejos y la de «extraterrestre» el conjunto de extraterrestres. Pero no considera que tales enunciados nos indiquen qué es la referencia. Para él, **averiguar qué es la referencia, es decir.Cuál es la naturaleza de la correspondencia entre las palabras y las cosas**, es un problema apremiante...para mi, hay poco que decir sobre lo que es la referencia dentro de un esquema conceptual, aparte de estas tautologías. **La idea de que es necesaria una conexión causal es refutada por el hecho de que el «extraterrestre» se refiere sin duda a extraterrestres, hayamos interactuado causalmente con extraterrestres o no.** ... El externalista renuncia de este modo al requisito de que tengamos alguna conexión real (por ejemplo, una conexión causal) con todas las cosas a las que podemos referirnos, exigiendo sólo que los términos básicos se refieran a géneros de cosas (y relaciones) con los que tenemos alguna conexión real. El externalista afirma que utilizando los términos básicos en combinaciones complejas podemos construir expresiones descriptivas que se refieran a géneros de cosas con las que no tenemos ninguna conexión real, y que pudieran incluso no existir (por ejemplo, extraterrestre). Cuando usamos la palabra «caballo» no sólo nos referimos a todas las demás cosas del mismo tipo.

[El caso del poeta que hablaba de «nenúfares» sin haberlos visto nunca y se sorprende por su belleza al conocerlos en la vida real o del señor que cuando habla de «campo» se refiere a ellos como el sitio donde los pollos se pasean vivos].VA

No obstante, debemos observar en este punto que «del mismo tipo que» es una expresión que no tiene sentido si no es desde un sistema categorial que señale qué propiedades no cuentan como semejanzas. Después de todo, cualquier cosa es del mismo tipo que cualquier otra de varias maneras. Todo este alambicado relato cerca de cómo nos referimos a algunas cosas en virtud del hecho de que estén conectadas con nosotros por cadenas causales del tipo apropiado y que en virtud del hecho de que son «del mismo tipo que» las cosas conectadas con nosotros por cadenas causales del tipo apropiado, e incluso a otras por «descripción», no es tan falso como ocioso. Lo que hace que los caballos con los que no he interactuado sean «del mismo tipo» que aquellos con los que sí lo he hecho, es que tanto los primeros como los últimos son caballos. La formulación del problema por parte del realista metafísico lo disfraza una vez más, como si hubiera que empezar con todos esos objetos en sí mismos, adquiriendo entonces un tipo peculiar de lazo con algunos pocos (los caballos con los que tengo conexión real, vía cadena causal del tipo apropiado), topándonos entonces con el problema de conseguir que mi palabra («caballo») cubra no sólo aquellos objetos que he «enlazado», sino también aquellos otros que no puedo enlazar... Y la «solución» a este pseudo-problema, tal como yo lo considero, —la «solución» del realista metafísico— es decir que la palabra cubre *automáticamente* no sólo los objetos que he enlazado, sino también los objetos que son del mismo tipo— en sí mismos. Pero entonces lo que se afirma, después de todo, es que el mundo consta de objetos que se Auto-identifican, pues afirmar que es el mundo, y no los sujetos pensantes, el que clasifica las cosas en géneros, significa precisamente eso.

**PUTNAM, H.** (1988:45-**48**). *Razón, verdad e historia*. Madrid: Tecnos.

Yo diría que el mundo *sí* consiste en Objetos que se Auto-identifican en un sentido —pero en un sentido no asequible al externalista. Si, como mantengo, **los propios objetos son tanto contruidos como descubiertos**, son tanto producto de nuestra invención conceptual como del factor «objetivo» de la experiencia, el factor independientemente de nuestra voluntad, entonces los objetos pertenecen intrínsecamente a ciertas etiquetas; porque esas etiquetas son las herramientas que usamos para construir una versión del mundo en la que tales objetos ocupan un lugar preeminente. Pero este tipo de objetos que se Auto-Identifican no es independiente de la mente; y lo que el externalista quiere es concebir el mundo como si consistiese de objetos que son independientes de la mente y que al mismo tiempo de Auto-Identifican. Y esto es lo que no se puede hacer.



**PUTNAM, H.** (1988:**48**). *Razón, verdad e historia*. Madrid: Tecnos.

## **INTERNALISMO Y RELATIVISMO**

El internalismo no es un fácil relativismo que afirma que «todo vale». Negar que tenga sentido preguntar si nuestros conceptos «se emparejan» con algo completamente incontaminado por la conceptualización es una cosa. Pero inferir a partir de esto que cualquier esquema conceptual es tan bueno como cualquier otro sería otra muy distinta... El internalismo no niega que haya *inputs experienciales* en el conocimiento; el conocimiento no es un relato que tenga otra constricción que la coherencia interna; lo que niega es que existan *inputs* que no estén configurados en alguna medida por nuestros conceptos, por el vocabulario que utilizamos para dar cuenta de ellos y para describirlos, o inputs que admitan una sola descripción, independiente de toda opción conceptual. Hasta la descripción de nuestras propias sensaciones, tan estimada como punto de partida del conocimiento por generaciones de epistemólogos, está profundamente afectada (como lo están las mismas sensaciones, dicho sea de paso) por multitud de opciones conceptuales. Los propios *inputs* sobre los que se basa nuestro conocimiento están conceptualmente contaminados. Pero es mejor tener *inputs* contaminados que no tener *inputs* de ninguna clase. Y si todo lo que tenemos son *inputs* contaminados, aun así no tenemos poco.

**PUTNAM, H.** (1988:49-50). *Razón, verdad e historia*. Madrid: Tecnos.

Lo que hace que un enunciado, o sistema completo de enunciados —una teoría o esquema conceptual— sea racionalmente aceptable es, en buena parte, su coherencia y ajuste; la coherencia de las ciencias «teóricas» —o menos experienciales— entre sí y con las creencias más experienciales; y también la coherencia de las creencias experienciales con las teóricas. Según a teoría que voy a desarrollar, nuestras concepciones de coherencia y aceptabilidad están profundamente entretajadas en nuestra psicología. Dependen de nuestra biología y de nuestra cultura y no están, en absoluto, «libres de valores». Pero son nuestras concepciones, y lo son de algo real. Definen un tipo de objetividad, objetividad para nosotros, si bien ésta no es la objetividad metafísica del punto de vista del Ojo de Dios. Objetividad y racionalidad humana es lo que tenemos; y tener esto es mejor que no tener nada.

Rechazar la idea de una perspectiva «externa» coherente, una teoría que simplemente es verdadera en sí misma, dejando a un lado todo posible observador, no es *identificar* la verdad con aceptabilidad racional. La verdad no puede ser tan sólo aceptabilidad racional por una razón fundamental; se supone que la verdad es una propiedad perenne de un enunciado, mientras que la justificación puede perderse. Con toda probabilidad, en enunciado «La tierra es plana» era racionalmente aceptable hace 3.000 años, pero no lo es hoy. No obstante sería erróneo decir que «La tierra es plana» era verdadero hace 3.000 años, ya que ellos significaría que la tierra ha cambiado de forma. En realidad, la aceptabilidad racional es relativa tanto a un tiempo como a una persona. Es además una cuestión de grado; a veces la verdad es presentada como una cuestión de grado ... de precisión el enunciado, y no de su grado de aceptabilidad o justificación.

**PUTNAM, H.** (1988:50). *Razón, verdad e historia*. Madrid: Tecnos.

En mi opinión, esto no muestra que la perspectiva externalista sea al fin y al cabo correcta, sino que la verdad es una *idealización* de la aceptabilidad racional. ..

Quizá parezca que explicar la verdad en términos de justificación bajo condiciones ideales es explicar una noción clara en términos de otra vaga. Pero la «verdad» no es tan clara cuando nos alejamos de ejemplos tan banales como «La nieve es blanca». Y, en cualquier caso, no estoy intentando ofrecer una definición *formal*, sino una elucidación informal, de la noción de verdad.

... las dos ideas clave de la teoría de la verdad-idealización son las siguientes: (1) la verdad es independiente de la justificación aquí y ahora, no independiente de toda justificación. Afirmar que un enunciado es verdadero es afirmar que podría ser justificado. (2) Es de esperar que la verdad sea estable o «convergente»; si tanto en un enunciado como su negación pueden ser «justificados», no tiene sentido pensar que tal enunciado posee un valor de verdad, por mucho que las condiciones fueran tan ideales como uno soñase alcanzar.

**PUTNAM, H.** (1988:50-51). *Razón, verdad e historia*. Madrid: Tecnos.

## **LA TEORÍA DE LA «SIMILITUD»**

La teoría que afirma que **la verdad es una correspondencia** es la más natural, desde luego. Quizá sea imposible hallar algún filósofo anterior a Kant que no mantuviera una teoría de la verdad-correspondencia.

Michael Dummett ha trazado recientemente [*What is a Theory of Meaning I, II en Truth and Other Enigmas. Harvard, 1980*] una distinción entre las perspectivas no-realistas (es decir, las que estoy denominando «internalistas») y las **reduccionistas**, con vistas a señalar que los reduccionistas pueden ser realistas metafísicos, es decir, suscriptores de la teoría de la verdad-correspondencia. El reduccionismo con respecto a una clase de afirmaciones (por ejemplo, afirmaciones con respecto a eventos mentales), es la concepción que mantiene que las afirmaciones de esa clase son «hechas verdaderas» por hechos que se encuentran fuera de esa clase. Por ejemplo, **de acuerdo con cierto tipo de reduccionismo, los hechos acerca de la conducta «hacen verdaderas» las afirmaciones acerca de eventos mentales.** Por poner otro ejemplo, la opinión del **obispo Berkley de que todo lo que «realmente existe» son las mentes y las sensaciones también es reduccionista**, puesto que sostiene que las oraciones acerca de mesas, sillas y otros «objetos materiales» ordinarios son hechas efectivamente verdaderas por hechos acerca de sensaciones.

**PUTNAM, H.** (1988:51). *Razón, verdad e historia*. Madrid: Tecnos.

Si un punto de vista es reduccionista con respecto a las afirmaciones de un tipo, pero con el único ánimo de insistir en la teoría de la verdad-correspondencia para las oraciones de la clase *reductora*, entonces el punto de vista es realista metafísico en su raíz.. Un punto de vista auténticamente no-realista es no-realista en todo su recorrido. Es frecuente cometer el error de considerar a los filósofos reduccionistas como no-realistas, pero Dummett sin duda está en lo cierto; ...

La forma más vetusta de la teoría de la verdad-correspondencia, que perduró aproximadamente 2.000 años, es la que los filósofos antiguos y medievales atribuyeron a Aristóteles. No estoy seguro de que Aristóteles realmente la mantuviese, aunque el lenguaje que emplea así lo sugiere. La denominare teoría de **referencia-similitud**, ya que sostiene que la relación que se da entre nuestras representaciones mentales y los objetos externos a los que se refieren es la relación de **similitud** literal.

La teoría empleó, como las teorías modernas, la idea de la representación mental. Esta representación, imagen mental del objeto externo, fue llamada *fantasma* por Aristóteles. La relación habida entre el fantasma y el objeto externo son similares (comparten la forma), la mente, el tener acceso al fantasma, también accede directamente a la propia forma del objeto.

El propio Aristóteles dice que el fantasma no *comparte* con el objeto propiedades, como la rojez (es decir, la rojez en nuestras mentes no es la misma propiedad literal que la rojez del objeto), que pueden ser percibidas por un solo sentido, pero si comparte propiedades como la longitud o la forma que pueden ser percibidas por más de un sentido (que son las sensibles comunes, en oposición a los «sensibles propios»).

**PUTNAM, H.** (1988:53-54). *Razón, verdad e historia*. Madrid: Tecnos.

## **LA TEORÍA KANTIANA DEL CONOCIMIENTO Y DE LA VERDAD**

Sugiero que la mejor lectura de **Kant** consiste en considerarlo como el primer autor que propuso lo que he denominado la **perspectiva «internalista»** o «realista interna» con respecto a la verdad, pese a que Kant en ningún momento afirme explícitamente que lo esté haciendo.

Para empezar, no hay duda de que Kant consideró inaceptables tanto el idealismo subjetivo de Berkeley -[*mentes y sensaciones*]... como el idealismo causal —el punto de vista que pretende que sólo percibimos directamente sensaciones e inferimos los objetos materiales mediante algún tipo de problemática inferencia—. Según Kant, la opinión según la cual es únicamente una hipótesis muy dudosa que hay una mesa delante de mí mientras escribo estas páginas es una provocación escandalosa.

En segundo lugar, creo que **Kant observó con claridad cómo operaba el argumento de Berkeley**: cómo **dependía de la teoría de la referencia-similitud**, y cómo la refutación de esa teoría era un requisito para la del argumento de Berkeley. Atribuyo aquí a Kant un punto de vista que Kant no expresa con estas palabras (en realidad, sólo recientemente **se llama «referencia» a la relación entre los signos mentales y aquello que representan**, si bien tal relación constituye un antiguo problema).

Sugiero que (como primera aproximación) la forma óptima de leer a Kant es como si generalizase a todas las cualidades lo que Locke afirmó con respecto a las cualidades primarias: a las simples, a las primarias y, del mismo modo, a las secundarias...

**PUTNAM, H.** (1988:54-55). *Razón, verdad e historia*. Madrid: Tecnos.

¿Qué se sigue de que todas las propiedades sean secundarias? Se sigue que todo lo que decimos sobre un objeto tiene la forma: éste es tal como nos afecta de tal y cual modo. **Nada de lo que afirmamos acerca de un objeto describe el objeto tal como es «en sí mismo», independientemente de su efecto sobre nosotros**, sobre seres con naturaleza racional y con nuestra constitución biológica. Se sigue también que no podemos suponer ningún tipo de similitud («similitude» en inglés de Locke) entre nuestra idea de un objeto y cualquiera que sea la realidad independiente-de-la-mente responsable de nuestra experiencia del objeto. **Nuestras ideas de los objetos no son copias de cosas independientes de la mente.**

Así es como Kant describe gran parte de la situación. Kant no duda de que hay alguna realidad independiente de la mente; para él es virtualmente un postulado de la razón. Alude con diversos términos a los elementos de esta realidad independiente de la mente: cosas-en-sí-mismas (Ding an sich); objetos nouménicos o noumeno; colectivamente, el *mundo nouménico*. Pero no podemos formarnos una concepción real de tales cosas nouménicas; la noción del mundo nouménico es más bien una límite del pensamiento ... que un concepto claro. Esta noción se concibe hoy como un elemento metafísico innecesario en el pensamiento de Kant. (Pero quizá Kant esté en lo cierto, quizá no podamos dejar de pensar que hay, de algún modo, una «base» independiente de la mente para nuestra experiencia, aun cuando las tentativa de hablar de ella nos conduzcan de inmediato al sin sentido).

**PUTNAM, H.** (1988:55). *Razón, verdad e historia*. Madrid: Tecnos.

**Al mismo tiempo, el discurso sobre «objetos empíricos» ordinarios no trata acerca de cosas-en-si-mismas, sino acerca de cosas-para-nosotros.**

El punto realmente sutil es que Kant considera que estos argumentos se aplican tanto a los objetos externos como a las sensaciones (objetos del sentido externo). Esto puede parecer extraño. ¿Acaso hay algún problema en que una idea se corresponda o no a una sensación?, pero Kant sigue una pista interesante y profunda.

Supongamos que experimento la sensación E, y que la describo afirmando que **L**” es la sensación de «rojo», por ejemplo. Si «rojo» significa meramente como esto, la afirmación completa únicamente significa «E es como esto» (señalando a E), es decir, E es como E —y al fin y al cabo, no he realizado ningún juicio. Como dijo Wittgenstein, uno se degrada virtualmente a un gruñido. Por otra parte, si «rojo» es un genuino *clasificador* y estoy afirmando que esta sensación E pertenece a la misma clase que las sensaciones que llamo «rojas» en otras ocasiones, entonces mi juicio va más allá de lo inmediatamente dado, más allá de la pura *haecceitas* (Thatness), e involucra **una referencia implícita a otra sensaciones no experimentadas en el instante presente** y una referencia implícita al tiempo (el cual, de **acuerdo con Kant, no es algo nouménico, sino más bien una forma en la que ordenamos las «cosas-para-nosotros»**).



**PUTNAM, H.** (1988:55-56). *Razón, verdad e historia*. Madrid: Tecnos.

¿Son las sensaciones que experimento en diferentes ocasiones, y que clasifico como *sensaciones* de rojo, «realmente» (nouménicamente) similares?

Esta pregunta carece de sentido. Si parecen (por ejemplo, si recuerdo sensaciones precedentes siendo semejantes a ésta, y anticipo que las futuras sensaciones que clasificaré de este modo se asemejarán a su vez a ésta, tal como en este momento la recuerdo) entonces **son similares-para-mi**.

Kant afirma una y otra vez, y de diferentes formas, que los objetos del sentido interno no son trascendentalmente reales (nouménicos) sino «trascendentalmente ideales» (**cosas-para-nosotros**) y que son directamente cognoscibles en el mismo grado en que pueden serlo los denominados objetos «externos». Las sensaciones a las que llamo «rojas» no pueden compararse directamente con objetos nouménicos en orden a observar si tienen la misma propiedad nouménica.

La razón por la que la afirmación «Todas las propiedades son secundarias» constituye sólo una *primera aproximación* al punto de vista de Kant es ésta: la afirmación «**Todas las propiedades son secundarias**» (es decir, todas las propiedades son facultades), sugiere que decir que una silla está hecha de pino, o de cualquier otra cosa, **es atribuible una facultad (la disposición que a nosotros nos parezca hecha de pino)** a un objeto nouménico; decir que la silla es marrón es atribuir una facultad diferente a ese mismo objeto nouménico, y así sucesivamente. Según este punto de vista, habría un único objeto nouménico correspondiente a cada objeto de lo que Kant llama «la representación», es decir, un único objeto nouménico correspondiente a cada cosa-para-nosotros. Pero **Kant niega explícitamente esto. Y este es el punto donde casi afirma que está abandonando la teoría de la verdad-correspondencia.**

**PUTNAM, H.** (1988:56-57). *Razón, verdad e historia*. Madrid: Tecnos.

En realidad, **Kant** no afirma que está abandonando la teoría de la verdad-correspondencia. Por el contrario, **afirma que la verdad es «la correspondencia de un juicio con su objeto»**. Pero esto es lo que **Kant llamó «definición nominal de verdad»**. En mi opinión, sería un grave error identificarla con lo que el realista metafísico quiere decir con «teoría de la verdad-correspondencia». ... hemos de ver si mantuvo una concepción realista acerca de lo que llamó «el objeto» de un juicio empírico.

En opinión de Kant, cualquier juicio acerca de objetos externos o internos (cosas físicas o entidades mentales) afirma que el mundo nouménico, como un todo, es de tal modo que ésta es la descripción que constituiría un ser racional (un ser con nuestra naturaleza racional), dada la información de que dispone un ser con nuestros órganos sensoriales (un ser con nuestra naturaleza sensible). En ese sentido, el juicio adscribe una facultad. Pero la facultad se adscribe al mundo nouménico como un todo; no debemos pensar que ya en nuestra representación hay sillas, caballos y sensaciones, habrá sillas nouménicas, caballos nouménicos y sensaciones nouménicas que les correspondan. **No hay una correspondencia uno-a-uno entre las cosa-para-nosotros y las cosas en sí mismas**. Kant no sólo renuncia a la idea de similitud entre nuestras ideas y las cosas en sí mismas; renuncia incluso a la idea de isomorfismo abstracto. Y esto significa que en su filosofía no hay una teoría de la verdad-correspondencia.

**PUTNAM, H.** (1988:57). *Razón, verdad e historia*. Madrid: Tecnos.

**Entonces, ¿qué es un juicio verdadero?** Kant cree que tenemos conocimiento objetivo: conocemos leyes de las matemáticas, leyes de la geometría, leyes de la física y muchos enunciados sobre objetos individuales —objetos empíricos, cosas para nosotros. El uso del término «conocimiento» y el uso del término «objetivo» comporta la afirmación de que a pesar de todo, hay una noción de verdad. **Pero si no es la correspondencia con la forma en que las cosas son en sí mismas, ¿qué es la verdad?**

La única respuesta que se puede extraer de los escritos de Kant, es, como he dicho, ésta: un fragmento de conocimiento (es decir, un «enunciado verdadero») es un enunciado que aceptaría un ser racional, a partir de una cantidad suficiente de experiencia de la clase que los seres con nuestra naturaleza pueden obtener efectivamente. Ni tenemos acceso ni podemos concebir la «verdad» en cualquier otro sentido. **La verdad es bondad última de ajuste.**

**PUTNAM, H.** (1988:57-58). *Razón, verdad e historia*. Madrid: Tecnos.

### **LA ALTERNATIVA EMPIRISTA**

Pese al punto al que ha llegado nuestro argumento, aún sería posible que un filósofo evitase el abandono de la teoría de la verdad-correspondencia y de la teoría de **la referencia-similitud** restringiéndolas a *sensaciones e imágenes*. Muchos filósofos continúan creyendo, **incluso después de Kant**, que la similitud **es el mecanismo por el que somos capaces de referirnos a nuestras propias sensaciones ... y que éste es el caso principal de referencia**, desde un punto de vista epistemológico.

De cara a observar por qué este recurso no resuelve el problema, recordemos que el núcleo del argumento de Berkeley era su aseveración de que **nada puede parecerse a una «idea» excepto otra «idea»**, es decir, que **no puede haber semejanza entre lo mental y lo físico**. De acuerdo con Berkeley nuestras ideas pueden parecerse a otras entidades mentales, pero no pueden parecerse a la «materia».

Debemos detenernos en esta afirmación y percatarnos de que es falsa, y en un aspecto importante. De hecho, cualquier cosa es similar a cualquier otra en infinitamente muchos aspectos. Mi sensación de la máquina de escribir en este instante y la moneda en mi bolsillo son ambas similares en cuanto que algunas de sus propiedades (que la sensación se produzca ahora mismo y que la moneda esté ahora mismo en mi bolsillo) son efectos de mis acciones pretéritas; y no estaría teniendo la sensación si no me hubiese sentado a escribir a máquina; y la moneda no estaría en mi bolsillo si no la hubiese metido ahí. Tanto la sensación como la moneda existen ambas en el siglo XX. He descrito en castellano tanto la sensación como la moneda, etc. Solo la ingenuidad y el tiempo limitan el número de semejanzas que uno puede encontrar entre cualesquiera objetos.

**PUTNAM, H.** (1988:58-59). *Razón, verdad e historia*. Madrid: Tecnos.

«Similitud» puede tener un significado más restringido en un contexto determinado, por supuesto. ... De este simple hecho se sigue ya que la idea de que la similitud es un mecanismo **privado de referencia** debe conducir a un regreso infinito.

De aquí no se concluye que no existen términos que tengan su lógica adscrita por medio de una teoría de la similitud, ni tampoco que no existan términos que se refieren a cosas que estén conectadas con nosotros por determinados tipos de cadenas causales. La conclusión es que la ni la similitud ni la conexión causal pueden ser los mecanismos de la referencia únicos y fundamentales\*

[VA: “Qué manera más curiosa, de recordar tiene uno, qué manera más curiosa/ hoy recuerdo mariposas que ayer fueron solo humo” Silvio Rodríguez]

**PUTNAM, H.** (1988:61). *Razón, verdad e historia*. Madrid: Tecnos.

### **APREHENSIÓN DE «FORMAS» Y ASOCIACIÓN EMPÍRICA**

Un platónico o neoplatónico a la antigua usanza habría afrontado este problema de una manera mucho más simple. Tal filósofo habría dicho que cuando prestamos atención a una sensación particular también percibimos un Universal o una Forma, es decir, que la mente tiene la capacidad de aprehender propiedades en sí mismas, y no sólo de prestar atención a instancias de esas propiedades. Tal filósofo diría que es el nominalismo de Wittgenstein y Goodman, su negativa a mantener relaciones con las Formas y con la aprehensión directa de Formas, lo que les hace ver problemática la teoría de la similitud.

Aun cuando postular sin más una misteriosa facultad de «aprender Formas» difícilmente puede ser una solución, podría parecer que disponemos de una facultad análoga. **Las propiedades de las cosas toman parte en las explicaciones causales;** cuando experimento una sensación y ésta obtiene la respuesta «Esta es la sensación de rojo», mi respuesta está causada, en parte, por el hecho de que la sensación tiene una propiedad. Es cierto que **algunos filósofos son tan nominalistas que negarían por completo la existencia de entidades tales como las «propiedades»;** pero la propia ciencia no puede vacilar en hablar libremente de propiedades. Cuando el personaje de Wittgenstein (el hombre que quiere inventar un lenguaje privado) señala C y dice «E», **¿no podemos decir que lo que causó la respuesta «E» fue una interacción causal que involucra cierta propiedad, y que esta propiedad (sea cual sea) constituye la similitud relevante que otras sensaciones deben mantener con X para ser clasificados correctamente como «E»?**

**PUTNAM, H.** (1988:61-62). *Razón, verdad e historia*. Madrid: Tecnos.

**La observación de que es perfecta y científicamente legítimo hablar de propiedades es correcta:** pero **esto no ayuda a la rehabilitación del platonismo**. Interactuamos con propiedades sólo mediante la interacción con sus *instancias*; y estas instancias **son siempre instancias de muchas propiedades al mismo tiempo**. No hay tal cosa como una interacción precisa con una propiedad «en sí misma». El discurso sobre propiedades causalmente asociadas con una sensación no puede realizar la labor que la idea de Forma (única) de la sensación realizaba en la filosofía platónica.

Detallando: cuando experimento la sensación de azul experimento además una sensación de azul una sensación con la compleja propiedad de ser de tal suerte que debo clasificar en ese caso bajo esa particular etiqueta. Atendiendo meramente a esta sensación no aprehendemos una de estas propiedades. Discernir la propiedad asociada con mi sensación o con la etiqueta verbal precisamente en uno de estos modos conduce a nuestro viejo amigo, el problema de la cadena causal del tipo apropiado.

**PUTNAM, H.** (1988:**62-63**). *Razón, verdad e historia*. Madrid: Tecnos.

Para comprenderlo<sup>↑</sup>, observamos en primer lugar que cuando mi experiencia perceptiva total provoca la respuesta «Estoy experimentando la sensación de azul», no siempre estoy en lo cierto. Yo mismo he tenido la experiencia de referirme dos o tres veces a «el hombre del suéter azul», antes de que alguien me indicase que el suéter era verde. No quiero decir que el suéter pareciera azul; en cuanto la otra persona habló, me di cuenta de que estaba describiendo erróneamente el suéter...

... Por otra parte, tal como expuso Wittgenstein, según este criterio, cualquier cosa que me parezca ser correcta —es decir, se habrá abandonado la distinción entre elaborar un informe de mi sensación realmente correcto y elaborar un informe de mi sensación que me parezca correcto. Quizá deberíamos abandonarla, o al menos atenuarla; quizá— como Goodman parece sugerir— no tenga sentido que uno se pregunte si está realmente experimentando el tipo de sensación que piensa estar experimentando, salvo casos especiales — como el caso en que uno se retractaría de su informe si alguien le expresase sus dudas; **pero un realista metafísico no puede jugar la baza de abandonar esta distinción, ya que lo que caracteriza la postura del realismo metafísico es la tajante distinción entre lo que realmente es y lo que uno juzga que es.**



**PUTNAM, H.** (1988:63-64). *Razón, verdad e historia*. Madrid: Tecnos.

**¿SE PUEDE ESTAR SIEMPRE EQUIVOCADO CON RESPECTO A LA CUALIDAD DE LAS SENSACIONES QUE SE TUVIERON EN EL PASADO?**

... según la teoría de la *similitud*, la respuesta es siempre «sí», porque de acuerdo con esta teoría, mis sensaciones previas o bien son o bien no son similares a las sensaciones que ahora describo mediante diversas etiquetas verbales, como «sensación de rojo», «dolor», etc., y si lo son o no es cuestión completamente distinta de si las clasifique entonces bajo esas mismas etiquetas verbales. A lo mejor el mundo es de tal forma que lo que llamamos «sensación de rojo» en un minuto par a partir del comienzo de la crisis cristiana es cualitativamente similar a lo que llamamos «sensación verde» en un minuto impar, pero **nuestra memoria siempre nos engaña, de tal manera que nunca lo notamos**. La sensación que clasifiqué bajo la etiqueta verbal «sensación de rojo» hace un minuto no sería similar a la sensación que ahora clasifico bajo la misma etiqueta.

... Si suponemos que la noción de similitud se autointerpreta, este caso podría redesccribirse como un caso en el que la relación que el observador externo —que nos está hablando del caso— llama «similitud» simplemente difiere de la relación que nosotros llamamos «similitud». **Si adoptamos este punto de vista, fracasará la hipótesis de que estamos realmente equivocados con respecto a nuestras sensaciones pasadas**: desde un punto de vista internalista no existe ninguna noción inteligible de similitud de sensaciones en diferentes tiempos, aparte de nuestros estándares de aceptabilidad racional.

**PUTNAM, H.** (1988:64). *Razón, verdad e historia*. Madrid: Tecnos.

## **OTRA VEZ LA TEORÍA DE LA VERDAD-CORRESPONDENCIA**

El lector puede haberse convencido ya de que la teoría de la referencia-similitud esta definitivamente muerta. Pero ¿por qué hemos de concluir que debe abandonarse la teoría de la verdad-correspondencia? Aun cuando la noción de «similitud» entre nuestros conceptos y sus referentes no solucione el problema, ¿no podría haber algún tipo de isomorfismo abstracto, o sin ser literalmente isomorfismo, **algún tipo de proyección de conceptos sobre cosas en el mundo (independiente de la mente)? ¿No podría definirse la verdad en términos de tal isomorfismo o «proyección»?**

Esta sugerencia tropieza con la dificultad, **no de que no existan correspondencias entre palabras o conceptos y otras entidades, sino de que existen demasiadas correspondencias.** Para seleccionar exactamente una correspondencia entre las palabras o los signos mentales y las cosas independientes de la mente, deberíamos tener de antemano acceso referencial a las cosas independientes de la mente. No podemos seleccionar una correspondencia entre dos cosas tomando como fulcro a una de ellas (o haciendo cualquier otra cosa sólo a una de ellas); usted no puede seleccionar una correspondencia entre conceptos y supuestos objetos nouménicos sin tener acceso a estos últimos.

**PUTNAM, H.** (1988:65). *Razón, verdad e historia*. Madrid: Tecnos.

Una vía para entender esto es la que sigue. Las teorías incompatibles pueden ser efectivamente intertraducibles en algunas cuestiones... La teoría del campo de Maxwell y la teoría de los potenciales retardados son incompatibles desde un punto de vista metafísico, ya que o bien hay o bien no hay agentes causales (los «campos») que medien la acción de partículas separadas entre sí (diría un realista). Pero las dos teorías son matemáticamente intraducibles. Así que, si hay una «correspondencia» hacia cosas nouménicas que haga verdadera a una de estas teorías, entonces se puede definir otra correspondencia que haga lo propio con la otra.

Si se considera que todo lo que hace verdadera a una teoría es una correspondencia abstracta (no importa cuál), entonces las teorías incompatibles pueden ser verdaderas.

Para un internalista, esto no es objetable: ¿por qué no han de existir a veces esquemas conceptuales igualmente coherentes, pero incompatibles, que se ajusten igualmente bien con nuestras creencias experienciales? **Si la verdad no es una (única) correspondencia, entonces se abre la posibilidad de cierto pluralismo. Pero la motivación del realista metafísico es salvar la noción del Punto de Vista del Ojo de Dios, esto es, La Teoría Verdadera.**

**PUTNAM, H.** (1988:65-66). *Razón, verdad e historia*. Madrid: Tecnos.

**No sólo puede existir una correspondencia entre objetos y (lo que nosotros consideramos) teorías incompatibles** (es decir, los mismos objetos pueden ser lo que los lógicos llaman «un modelo» para teorías incompatibles), sino que aun cuando fijemos la teoría y los objetos, hay una cantidad infinita (si el número de objetos es infinito) de formas diferentes en las que pueden usarse los mismos objetos para construir un modelo para una teoría dada...

**Nos encontramos ante el óbito de una teoría que resistió alrededor de dos siglos.** El que persistiera tanto y de tantas formas, a pesar de sus obscuridades y contradicciones internas, presentes desde un principio, **da fe de la naturalidad y de la fuerza del punto de vista del Ojo de Divino.** Kant fue el primero en enseñarnos la insatisfacibilidad de tal deseo, pensó que, no obstante, formaba parte de nuestra propia naturaleza racional (sugirió sublimar este «impulso totalizador» en el proyecto de realizar el mayor bien en el mundo, reconciliando los órdenes moral y empírico en un sistema de instituciones sociales y relaciones individuales perfeccionado). **La continua presencia de este impulso natural pero insatisfacible quizá sea el motivo de fondo de los falsos monismos y dualismos que proliferan en nuestra cultura; sea como sea, nos hemos quedado sin el punto de vista.**

**PUTNAM, H.** (1988:68-69). *Razón, verdad e historia*. Madrid: Tecnos.

### **MENTE Y CUERPO-PARALELISMO, INTERACCIONISMO E IDENTIDAD**

Fueron varios los autores que propusieron la primera formulación contemporánea de esta teoría de la **identidad**, siendo uno de los más conocidos el filósofo australiano J.J.C. Smart. En un principio se sugirió que una sensación, por ejemplo, **la sensación particular de azul es idéntica a cierto estado neurofisiológico**. Creo que fui yo quien sugerí por primera vez el punto de vista denominado **funcionalismo**, que **es una variante de la anterior afirmación**. Esta concepción admite la existencia de **una identidad en el lugar señalado**, pero niega que en el otro término de la identidad figure el tipo de propiedad que Smart señala. **De acuerdo con el funcionalista, el cerebro tiene propiedades que, hasta cierto punto, son no-físicas**.

Ahora bien, ¿qué quiero decir al afirmar que el cerebro tiene propiedades no físicas? **Quiero decir que posee propiedades que son definibles en términos que no aluden a la física o a la química cerebral**. Si nos parece extraño que un sistema físico tenga propiedades que no son físicas, prestamos atención a un ordenador. Un ordenador tiene muchas propiedades físicas. Tiene cierto peso, por ejemplo, tiene también cierto número de circuitos o *chips*, o cualquier otra cosa. Tiene propiedades económicas... y propiedades funcionales, tales como tener un programa... propiedad no-física... Un espíritu incorpóreo podría presentar cierto programa, un cerebro podría presentar cierto programa, un cerebro podría presentar cierto programa, una máquina podría presentar cierto programa, pero la organización funcional de los tres (el espíritu incorpóreo, el cerebro y la máquina) podría ser exactamente en la misma aun cuando su materia, su sustancia, fuese completamente diferente.

**PUTNAM, H.** (1988:69). *Razón, verdad e historia*. Madrid: Tecnos.

Las propiedades psicológicas presentan la misma característica; la misma propiedad psicológica (por ejemplo, estar iracundo) puede ser una propiedad perteneciente a miembros de miles de especies diferentes, que pueden poseer una composición física o química completamente distinta (algunas especies podrían ser extraterrestres; y quizá los robots muestren ira algún día). **La sugerencia del funcionalista es que la teoría «monista» más plausible que se pueda defender en el siglo XX, la teoría que evita tratar a la mente y a la materia como dos tipos separados de sustancias o como dos reinos separados de propiedades, es la que identifica propiedades psicológicas y propiedades funcionales.**

**Todavía hoy me inclino a pensar que esa teoría es correcta;** o al menos que constituye una descripción *naturalista* correcta de la relación mente-cuerpo. Las descripciones «mentalistas» de esta relación también son correctas, aunque no son reducibles a la imagen del mundo que llamamos «Naturaleza» **(en realidad, las nociones de «racionalidad», «verdad» y «referencia» pertenecen a esta versión mentalista).** ... este hecho no me desanima: como ha puesto de relieve Nelson Goodman, uno de los atractivos del no-realismo es que permite la posibilidad de versiones correctas del mundo alternativas. A pesar de todo, **me atrae la idea de que una de las versiones correctas del mundo es la naturalista,** en la que las formas-de-pensamiento, imágenes, sensaciones, etc., son acontecimientos físico funcionalmente caracterizados; lo que deseo discutir aquí es **una dificultad de la teoría funcionalista,** que se me planteó hace ya varios años: **la teoría funcionalista halla dificultades con el carácter cualitativo de las sensaciones.**

**PUTNAM, H.** (1988:69-70). *Razón, verdad e historia*. Madrid: Tecnos.

Cuando se piensa que estados psicológicos puros y relativamente abstractos, por ejemplo, en lo que he denominado creencias «puestas entre paréntesis», esto es, un pensamiento considerado sólo en su contenido «nocional», o en **estados emocionales tan difusos como estar celoso o iracundo**, su identificación con estados funcionales de todo el sistema parece sumamente plausible; pero cuando se piensa en experimentar una cualidad dada, por ejemplo, experimentar una tonalidad particular de azul, la identificación ya no es plausible.

Durante muchos años he utilizado en mis clases una variante del famoso ejemplo del «**espectro invertido**». El ejemplo del espectro invertido (que aparece en los escritos de Locke), atañe a un tipo que ve las cosas de tal modo que el azul le parece rojo y el rojo azul (o del tal modo que sus colores subjetivos parecen estar en negativo y no en positivo). La primera reacción al oír un caso así podría ser la siguiente: «Pobre tipo, es digno de lástima». Pero ¿cómo podría saberlo algún otro? Cuando ve algo azul le parece rojo, pero desde la infancia se le ha enseñado a llamar «azul» a ese color, de modo que si se le pregunta cuál es el color del objeto, contestaría que azul. De forma que nadie lo sabría nunca.

**PUTNAM, H.** (1988:70). *Razón, verdad e historia*. Madrid: Tecnos.

Supongamos ahora que adoptamos la siguiente teoría «funcionalista» con respecto al color subjetivo: «una sensación es la sensación de azul (es decir, tienen el carácter *cualitativo* que describo de esa forma en este momento) solo en el caso de que la sensación» (o el correspondiente acontecimiento físico en el cerebro) desempeñe el papel de señalar la presencia de azul objetivo en el entorno. Aunque esa teoría capta un sentido de la locución «sensación de azul», no es éste el sentido «cualitativo» deseado. **Si el rol funcional fuese idéntico al carácter cualitativo entonces no podría afirmarse que la cualidad de la sensación ha sufrido cambio.** (Si esto no se ve claro del todo, imaginemos que una vez ocurrida la inversión del espectro, y después de haber aprendido a compensarla lingüísticamente, usted experimenta un ataque de amnesia que le borra por completo de la memoria cómo solía ver los colores. En este caso, podría parecer que la sensación a la que usted llama ahora «sensación de azul» desempeña casi el mismo rol funcional que la sensación que usted solía llamar «**sensación de azul**», a pesar de que tiene un carácter distinto). Más la cualidad ha cambiado. **En este caso, la cualidad no parece ser un estado funcional.**

Si estos casos son realmente posibles, me parece que la baza más plausible puede jugar un funcionalista es decir: «Sí, pero el “carácter cualitativo” es precisamente la materialización física» y a continuación afirmar que para todo tipo especial de propiedad psicológica, para las cualidades, la formulación correcta de la teoría de la **identidad** es la más antigua. Si el lector(a) tiene inclinaciones honestamente **materialistas, probablemente piense que la propiedad de tener una sensación es una propiedad del cerebro.**



**PUTNAM, H.** (1988:70-71). *Razón, verdad e historia*. Madrid: Tecnos.

Y, probablemente, **la mayoría de las personas sostiene una de estas dos opiniones: la que afirma que los estados sensoriales están correlacionados con estados del cerebro o la que afirma que los estados sensoriales son idénticos a los estados del cerebro....** El debate cobra esta forma una y otra vez. He aquí como discurre habitualmente la discusión: dado que B está correlacionado con Q, ¿es B efectivamente idéntico a Q? Sabemos que este estado sensorial corre paralelo a este estado del cerebro: ¿es el primero idéntico al segundo o no lo es? A medida que la discusión va discurriendo por esos derroteros, **el concepto de correlación** va pareciendo menos problemático. La correlación no es una cuestión (muy) debatida, pues todo el mundo sabe que hay al menos una correlación. La identidad sí que se discute, porque es ahí donde radica el problema, claro está. Frente a esto, voy a intentar demostrar que hasta **la correlación es problemática... El problema no radicará en asumir el materialismo, sino en el hecho de que creemos que hay al menos una correlación.**

**PUTNAM, H.** (1988:71-72). *Razón, verdad e historia*. Madrid: Tecnos.

## **LA TEORÍA DE LOS A PRIORI**

Fue un cambio de clima epistemológico el que hizo posible el resurgimiento del interés hacia la **teoría de la identidad** y hacia otras teorías «monistas»; ... empezó a notarse en los comienzos de la década de los 60. La teoría de la identidad no se tomó en serio antes de esta década porque los filósofos «estaban seguros» de que era falsa, y estaban convencidos de que sabían que era falsa *a priori*, no sobre la base de la evidencia empírica (**¿qué tipo de evidencia empírica podría mostrar que un estado sensorial no puede ser un estado cerebral?**). Basta pensar sobre ello para comprender a priori que no tiene sentido decir que un estado sensorial es un estado cerebral, del mismo modo que no lo tiene decir que el número tres es azul. **Anteriormente a 1950 ó 1960, muchas personas estaban convencidas de que sabían, sin más, que los estados sensoriales no podían ser estados físicos.** Otras estaban convencidas de que sabían que las primeras estaban equivocadas. ... La mayoría estaba convencida de que los estados sensoriales no podían ser estados cerebrales, y una minoría estaba convencida de que la mayoría estaba equivocada. Cada parte sabía a priori que la otra estaba en un error.

En 1951, W.V. Quine publicó un artículo con el título «Dos dogmas del empirismo» ... Quine señaló cosas que creíamos saber a priori han tenido que ser revisadas. ... **Quine también mostró de forma convincente que las descripciones empiristas típicas de la noción de *a priori* —por ejemplo, la noción de «verdad por convicción»— eran incoherentes...**

**PUTNAM, H.** (1988:72-73). *Razón, verdad e historia*. Madrid: Tecnos.

Aunque Quine exagera las acusaciones contra la noción de *a priori*, tiene razón en un punto: nuestras nociones de racionalidad y de revisabilidad racional no se fijan mediante algún manual de reglas inmutable, ni están inscritas en nuestra naturaleza trascendental (como Kant pensaba) por excelente razón de que la propia idea de una naturaleza trascendental, una naturaleza que poseemos nouménicamente —sin tener en cuenta aquellos sistemas en los que podemos concebirnos histórica o biológicamente— carece de sentido. ... Con todo, esta revisión no puede ser ilimitada: de otro modo se disiparía el concepto de lo que podríamos llamar *racionalidad*; pero, por lo general, **no está a nuestro alcance establecer límites**. Dejando aparte los casos triviales (por ejemplo, «No todo enunciado es verdadero») **no podemos estar seguros de que en ningún contexto sería racional desechar un enunciado que se considera como una verdad necesaria** (y de modo legítimo, *en un contexto determinado*). **Hemos de admitir que, en general, las consideraciones de simplicidad, de utilidad global y de plausibilidad pueden llevarnos a renunciar a algo que anteriormente tomábamos por un *a priori*, y que hacer esto es razonable**. La filosofía se ha convertido en anti-apriorística. Pero una vez hemos reconocido que la mayor parte de las verdades que considerábamos como verdades a priori tiene un carácter contextual y relativo, hemos renunciado también al único buen argumento que existía en contra de la identidad mente-cuerpo. Los teóricos de la identidad estaban obligados a poner de manifiesto este punto y obraron en consecuencia. Así que la situación cambió.

**PUTNAM, H.** (1988:**73**). *Razón, verdad e historia*. Madrid: Tecnos.

He estado haciendo uso de la noción de propiedad, aunque creo que hemos llegado a confundir al menos dos nociones de «propiedad» En relación con la noción más antigua solía emplearse **en término «predicado»** (por ejemplo, en la célebre pregunta «**¿Es la existencia un predicado?**»). La otra noción es la que usamos hoy al hablar de «**propiedades físicas**», «**magnitudes fundamentales**», etc. Cuando un filósofo tiene en mente la primera, considera frecuentemente que el discurso acerca de propiedades es **intercambiable con el discurso acerca de conceptos**. Según este filósofo, *Impropiedades no pueden ser idénticas a menos que constituya una verdad conceptual el que lo sean; en particular, la propiedad de experimentar una sensación con cierto carácter cualitativo no puede ser la misma propiedad que la de hallarse en cierto estado cerebral*, ya que los correspondientes predicados **no son sinónimos** (en el amplio sentido de «analíticamente equivalentes»), **y el principio de individuación para predicados consiste precisamente en que ser P es uno y el mismo predicado que ser Q si y solo si «es P» es sinónimo de «es Q»**.

**PUTNAM, H.** (1988:73-74). *Razón, verdad e historia*. Madrid: Tecnos.

No obstante, consideremos la situación que se crea cuando un científico afirma que la temperatura es energía cinético-molecular media. A primera vista, este enunciado identifica propiedades. **Lo que se afirma es que la *propiedad* de tener una temperatura particular es realmente (en algún sentido de «realmente») la misma propiedad que tener cierta energía molecular, o (de forma más general) que la *magnitud física* temperatura y la *magnitud física* energía cinético-molecular media son una y la misma propiedad.** Si esto último es correcto, entonces, ya que «X tiene tal y tal temperatura» y «X tiene tal y cual energía cinético-molecular media» no son oraciones *sinónimas*, aun cuando «tal y cual» sea el valor de la energía molecular correspondiente al valor «tal y tal» de la temperatura, **habremos de concluir que el físico llama «magnitud física» a alguna cosa distinta de lo que los filósofos han llamado «predicado» o «concepto».**

Siendo más explícito: la diferencia estriba en que mientras que para que los predicados P y Q sean los mismos se requiere la sinonimia de las expresiones «X es P» y «X es Q», esta sinonimia no es requisito para que la propiedad P y propiedad Q sean la «misma» propiedad. **Las propiedades, al contrario que los predicados, pueden ser «sintéticamente idénticas».**

PUTNAM, H. (1988:73-74). *Razón, verdad e historia*. Madrid: Tecnos.

**Si existe algo como la identidad sintética de propiedades, entonces ¿por qué no puede ocurrir que la propiedad de hallarse en cierto estado cerebral sea la misma propiedad que la de experimentar una sensación con cierto carácter cualitativo** (esto es algo que se encuentra muy en la línea de pensamiento de Spinoza), aun cuando ello constituya una verdad conceptual?... En resumen, nos encontramos con una ola de antiapriorismo y con un nuevo **sistema para la identidad *sintética de propiedades*, y gracias a ambas, los teóricos de la identidad, y en particular, el funcionalista, parecen entrar en acción automáticamente.**

**PUTNAM, H.** (1988:74). *Razón, verdad e historia*. Madrid: Tecnos.

## **CEREBROS ESCINDIDOS**

Ahora bien, quiero examinar lo que sucede cuando a estas dos cosas añadimos una tercera. ¿Qué sucede si un filósofo es (1) naturalista antiapriorista, (2) permite que haya algo como una identidad sintética de propiedades y (3) mantiene una concepción realista con respecto a la verdad, en la línea dura? Deseo afirmar que este mismo filósofo se verá enfrentado a serias dificultades epistemológicas.

Examinaremos un determinado tipo de experimento que los neurólogos han venido realizando en los últimos veinte años. Es el célebre experimento de los «cerebros escindidos» o experimento de disociación cerebral.

Quiero traer a colación la relevancia de este tipo de experimento con respecto a la teoría de la identidad y con respecto a lo que hasta este punto se ha dado por sentado en toda la discusión: **la idea de que hay una correlación.**

Según el modelo del cerebro como un **sistema cognitivo semejante a una computadora**, el cerebro posee **un lenguaje interno** (que podría ser innato, o una mezcla de «lenguaje» o **sistema de representación innato y lenguaje público**). Algunos filósofos han llegado incluso a inventar un nombre para este hipotético lenguaje cerebral, **el «mentalés»**. Consideremos lo que sucede cuando uno experimenta una sensación visual... la cosa podría suceder así: ↓

**PUTNAM, H.** (1988:74-75). *Razón, verdad e historia*. Madrid: Tecnos.

Cuando uno experimenta una sensación, se lleva a cabo un «juicio»: el cerebro tiene que «imprimir» algo así como «registrado el color rojo a las 12 en punto». De modo que la cualidad (llamémosle «Q») corresponde, entre otras cosas, a una grabación en mentalés. Del mismo modo se da un *input* hacia el centro de procesamiento verbal, conectado con las cuerdas vocales; estas últimas dan cuenta de la capacidad del cerebro para informar, el lenguaje público, «experimento rojo en este momento»... Estos «registros», «*inputs*» y demás acontecimientos pueden tener lugar en diferentes lóbulos del cerebro: si se secciona el *corpo calloso*, el lóbulo derecho de la persona (que no posee habla) puede ver (algo) rojo (o al menos dará una señal afirmativa a una pregunta escrita visible sólo para este lóbulo) pero si se le pregunta al sujeto «¿de qué color es la tarjeta?» responderá «No puedo ver la tarjeta». Y, finalmente, en algún punto se forman la huella o huellas de memoria (disociable en *memoria a corto plazo* y *memoria a largo plazo*). **Es casi seguro que no existe una cadena causal**; probablemente existen ramificaciones y reincorporaciones: es decir, una red causal.

**El problema consiste en que la psicología divide los eventos mentales de una forma excesivamente *discreta*. He aquí la sensación de azul: en este momento empieza, en este otro acaba. Pero las redes causales no son discretas.**

<b>No hay un único evento físico que sea el <i>correlato</i> de la sensación.</b>
---



**PUTNAM, H.** (1988:75). *Razón, verdad e historia*. Madrid: Tecnos.

**Si la teoría de la identidad está en lo cierto, el estado sensorial Q es idéntico a algún estado cerebral.** Un realista metafísico no puede considerar esta identidad como un asunto de convención o de decisión o como si se tuviese un componente convencional idéntico al estado cerebral Q. **La opinión es que, como cuestión de hecho, vivimos en un mundo en el cual lo que experimentamos como caracteres cualitativos de las sensaciones son en realidad las mismas propiedades que algunas de las propiedades que nos encontramos en otros campos como propiedades físicas de eventos cerebrales.** (O, mejor dicho, en los cuales la propiedad de experimentar una sensación con cierto carácter cualitativo es exacta y realmente la propiedad de hallarse en cierto estado cerebral).

... Supongamos también que cuando experimento esta cualidad de rojo, el estado sensorial en el que me hallo es idéntico a una disyunción de estados cerebrales. El estado sensorial no puede ser idéntico a un estado cerebral máximamente especificado, puesto que sabemos que seguiríamos teniendo la máxima experiencia aun cuando anulásemos una neurona, o cualquier elemento. Pero la propiedad puede ser disyuntiva, por ejemplo... o bien se están *descargando las neuronas pares en tal-y-tal* o bien se están *descargando aquellas cuyo ordinal es un número primo*. **Habría una descomunal colección de estados neurológicos tales que su disyunción constituiría la propiedad de experimentar algo rojo.**

**PUTNAM, H.** (1988:76-77). *Razón, verdad e historia*. Madrid: Tecnos.

Se puede eludir el compromiso con esa perspectiva. Por ejemplo, **Carnap habría dicho (al menos en cierto período) que el discurso acerca de objetos físicos es en realidad discurso acerca de sensaciones**, si bien muy derivado, y que la decisión de afirmar que un particular estado cerebral es idéntico a un estado sensorial Q es, en realidad, la decisión de modificar de cierta forma el lenguaje del discurso acerca de las propiedades físicas, de cambiar nuestro concepto de la propiedad física en cuestión.

Puesto que el discurso acerca de objetos y propiedades físicas es tan sólo una derivación del discurso acerca de sensaciones, podemos modificar las reglas. Pero este punto de vista no es el del realismo metafísico, al menos en lo que respecta a objetos materiales y propiedades físicas. Quien desee ser realista metafísico en lo que atañe a sensaciones, sin serlo con respecto a los *objetos materiales*, puede adoptar la teoría de la identidad (ya que considera al discurso acerca de los objetos materiales como flexible) afirmando simplemente «La adopto como un tipo de convención, como una **estipulación adicional de significado**». Ya que los significados no están fijados de antemano y ya que existía cierta textura abierta, desaparece el problema consistente en «¿Cómo podemos saber que el estado sensorial es idéntico a esta propiedad y no a alguna otra?». **Si lo que es esta propiedad es bastante vago, estamos autorizados a postular la identidad simplemente como una especificación de significado. Pero mi debate tiene como interlocutor a alguien que realmente piense que existe un mundo material ahí fuera, y que éste no es una mera derivación del discurso acerca de sensaciones;** alguien que realmente crea que hay propiedades físicas, y que sostenga que expresiones tales como «Se están descargando las neuronas en tal-y-tal canal» son predicados que definen nuestras propiedades físicas y que cualquiera de estas propiedades o bien es *idéntica* a este estado sensorial o bien no lo es.

**PUTNAM, H.** (1988:77). *Razón, verdad e historia*. Madrid: Tecnos.

De modo similar, creo que un filósofo como **Daniel Dunnett, quien piensa que el discurso acerca de sensaciones es sumamente vago, y no cree que haya una propiedad subjetiva bien definida que consista en hallarse en tal estado sensorial o en experimentar una sensación con tal carácter cualitativo**, podría adoptar una teoría de la identidad en la forma de estipulación de significado, aunque lo que se determinaría en esta ocasión sería el significado de los términos psicológicos y no de los términos de objetos físicos. Pero, una vez más, la actitud del realista metafísico radical no sería ésta.

Estoy pensando en un realista metafísico radical que discurra de esta guisa: «Sí, se en qué consiste esta propiedad psicológica (el estado sensorial). La he experimentado, puedo reconocerla. Creo que es la propiedad psicológica a la que me refiero. Se lo que son  $P_1$  y  $P_2$  y, por ende, lo que es  $(P_1 \text{ o } P_2)$ , y el estado sensorial o bien es idéntico a esta propiedad o bien no lo es», tal como un físico podría decir: «No hay ningún elemento convencional... en la decisión en favor de que la temperatura es energía cinético-molecular media, así que, o bien la temperatura es energía cinético-molecular media o bien es alguna otra propiedad».

**PUTNAM, H.** (1988:77-78). *Razón, verdad e historia*. Madrid: Tecnos.

El problema radica en que, si uno acepta el punto de vista del realista metafísico, existen muchas más posibilidades de las que la gente acostumbra a considerar. La primera posibilidad que se nos ocurre consiste en que el estado sensorial sea idéntico a la propiedad de que ocurran los eventos adecuados en el córtex visual y que se tenga adecuadamente grabado en «registro» en «mentales», y que se dé un *input* hacia el centro del habla y que se hallen formadas huellas de la memoria —es decir, el estado sensorial se concibe idéntico a la conjunción de todas las propiedades. Pero tan pronto consideramos la posibilidad de la disociación cerebral dejamos de estar seguros de que sea necesaria toda la conjunción. **Quizá la sensación sea precisamente cierto evento en el córtex visual (es decir, la propiedad de experimentar una sensación es «realmente» la propiedad de que en el córtex visual tenga lugar cierto evento).**

Supongamos que es así, de momento. Ahora bien, supongamos que podemos bloquear el proceso que produce el registro en mentales, o, al menos, que podemos bloquear el *input* que se dirige hacia el centro del habla. Imaginemos que le hemos mostrado al sujeto una tarjeta roja en el lado izquierdo de su campo visual (de forma que la tarjeta sólo es «visible para el lóbulo derecho», como dicen los neurólogos): El evento en cuestión, que tendría lugar en el córtex visual, se dará entonces en el lóbulo derecho; pero si le preguntamos al sujeto «¿Ha visto usted algo rojo?», contestaría «No».

... no nos queda más remedio que admitir que el sujeto no ha experimentado la sensación roja. Y, por lo tanto, tendremos que admitir la «refutación» de la teoría de que Q (el carácter cualitativo relevante) es idéntico a esa propiedad del córtex visual.↓

**PUTNAM, H.** (1988:78-79). *Razón, verdad e historia*. Madrid: Tecnos.

↑pero alguien podría objetar «No, de ningún modo ha refutado la teoría. Pues ¿qué tipo de observador nos ofrece? El sujeto tiene el cerebro partido por la mitad». En tanto dispongamos de un observador en condiciones normales, la propiedad Q es idéntica a esta propiedad del córtex visual...

La dificultad consiste en que existen teorías de la identidad *observacionalmente indistinguibles* con ello quiero decir que son teorías que conducen a las mismas predicciones con respecto a la experiencia de todos los observadores que se hallen en *condiciones normales*.

Consideremos la teoría que afirma que es imposible tener la sensación de rojo salvo que se dé el *input* en el centro del habla. ¿Cómo puede ser probada o refutada? Podríamos pensar que hay un modo, siempre que escindamos el cuerpo calloso y que parte de la memoria no atraviese la unidad de proceso verbal; a saber, primero le preguntamos al sujeto si experimenta la sensación de rojo. Este contesta: «No». A continuación volvemos a coser el cuerpo calloso ... y le preguntamos «¿Ha experimentado usted la sensación de rojo?». Podría contestar «Sí, pero esto es un disparate: usted sabe que la he experimentado y aun así me lo pregunta, y yo me he oído a mí mismo contestándole con toda sinceridad que no» (b). ¿Demostraría tal informe que *hubo* una sensación de rojo sin que se diese un *input* hacia el centro del habla?

No lo haría... Muchos neurólogos creen que el lóbulo derecho de los pacientes con el «cerebro escindido» es «consciente». En efecto, esto equivale a afirmar que a veces hay una sensación de rojo, a pesar de que no se dé ningún *input* hacia el centro del habla. Al menos un famoso neurólogo, **Eccles ... la conciencia es unitaria;** que el lóbulo disociado pueda «simular» una conducta consciente no demuestra que sea un segundo «*locus de conciencia*».

**PUTNAM, H.** (1988:79). *Razón, verdad e historia*. Madrid: Tecnos.

La teoría de los dos «loci» es más simple en un aspecto: afirma que ciertas capacidades conductuales (que posee el lóbulo derecho, aun cuando no posea habla) son suficientes para la conciencia, y esto casa con el hecho de que calificamos como conscientes a los animales (los cuales tampoco poseen habla). Pero existen muchas semejanzas entre un animal con el cerebro intacto, cuyos procesos cerebrales aún están integrados (aunque no involucre el habla), y una parte de un «cerebro escindido». Si el caso nos toca de cerca, si no tuviésemos una tendencia tan acusada hacia el realismo metafísico con respecto a las *sensaciones*, ¿no estaría más de acuerdo con nuestras intuiciones metodológicas considerarlo como un caso a legislar, en vez de una cuestión sobre la que disputar?

En suma, existen varias teorías sobre la identidad observacionamente indistinguibles. Si el teórico de la identidad está en lo cierto, da la impresión de que no hay forma humana de saber de que modo está en lo cierto, de saber cuál es el estado cerebral que es idéntico al estado sensorial (o está relacionado con éste).

**PUTNAM, H.** (1988:81-82). *Razón, verdad e historia*. Madrid: Tecnos.

Por último, quiero exponer tres categorías que con toda seguridad son falsas, pero cuyo rechazo es difícil o imposible en el caso de que el realismo metafísico sea correcto. Son estas: (1) **rojo<sub>h</sub> es**, después de todo, **un estado funcional** o casi-funcional, a saber, el estado consiste en **hallarse en cualquier estado material** (por ejemplo, físico) **que desempeñe en nuestras vidas con anterioridad el rol funcional de de señalar la presencia de rojo objetivo**. (2) **Las rocas tienen qualia** (es decir, en las rocas tienen lugar eventos cualitativamente similares a las *sensaciones visuales*). (3) **Las naciones son conscientes**.

... O, al menos, consideremos la hipótesis de que alguna de estas hipótesis es correcta (sin especificar cuál). Podríamos decir, «Bueno, pero estas hipótesis son disparates». El efecto, lo son. Pero todas y cada una conducen a la teoría «sana». Ninguna de ellas puede ser excluida sobre una base observacional o experimental, puesto que todas ellas son observacionalmente indistinguibles... Podríamos pensar que estas teorías pueden excluirse apelando al **principio metodológico: no se ha de atribuir una propiedad a un objeto sin ninguna razón...** ¿Qué ocurre con el argumento de que si nosotros tenemos qualia y el fisicalismo es verdadero ... entonces hay al menos un objeto físico en el cual tienen lugar eventos en todos los objetos físicos? Esta posibilidad se cerraría si pudiésemos mostrar que existe algo en la propia *cualidad* que le exige tener «rob» **funcional** particular que tiene para los seres humanos; pero los creyentes en los *qualia* como objetos metafísicamente reales nos dicen que es esto precisamente lo que no podemos hacer.

**PUTNAM, H.** (1988:**82-83**). *Razón, verdad e historia*. Madrid: Tecnos.

Consideremos la hipótesis de que el dolor es idéntico a un estado funcional adecuado, que puede manifestarse también en los organismos institucionales o en las naciones. En otras palabras, supongamos que cuando en España se declara «España está afligida por... » , realmente lo está. Nunca lo sabríamos, claro. Quizá el lector encuentre es este momento interesante y algo divertido el que un grupo pueda comportarse de manera semejante a como se comporta, cuando siente dolor, algo que realmente siente dolor; pero el lector no cree que España realmente sienta dolor. Según la hipótesis, el lector estaría equivocado: el espíritu nacional realmente estaría sufriendo dolor.

Esta hipótesis guarda relación con una interesante polémica en filosofía de la mente. A los **funcionalistas (entre los que me incluyo)** les gusta emplear el siguiente argumento «antichovinista»: las diferencias entre un robot y un ser humano, en cuanto a organización funcional, pueden reducirse, en principio, a pequeños detalles físico-químicos. Hasta podríamos encontrar un robot cuyo nivel funcional fuese correspondiente al nuestro. (E incluso podría tener un cuerpo «de carne y hueso», además de cerebro). Ahora bien, salvo que seamos unos «chovinistas de hidrógeno-carbono» y pensemos que el hidrógeno y el carbono son intrínsecamente más conscientes, ¿por qué no decir que este robot es una persona cuyo cerebro resulta tener más metal y menos hidrógeno y carbono?

Este argumento ha provocado la siguiente réplica... «Ahora bien», continúa la réplica, «no diríamos que tal cosa es “consciente”, ya que sabemos que, en realidad, consta sólo de seres minúsculos moviendo su cuerpo. Y esto muestra que una organización funcional adecuada (como la nuestra) **no basta para la aplicación de predicados como “consciente”**».



**PUTNAM, H.** (1988:**83**). *Razón, verdad e historia*. Madrid: Tecnos.

Una réplica a esta réplica... es negar que el «robot hidra-cefálico (así se le ha llamado) tenga la misma organización funcional que nosotros tenemos... Podría haber preguntado «¿Qué nos impide afirmar que el robot hidra-cefálico es consciente? Si el primer argumento es correcto (y creo que lo es), si el robot que posee un cerebro positrónico puede ser consciente, ¿por qué motivo el hecho de que las neuronas del robot hidra-cefálico» son más *conscientes* significa que ese ente, en su totalidad, es menos consciente? Después de todo, somos, hasta cierto punto, pequeños animales. Nuestras células son, hasta cierto punto, pequeños animales individuales. Y quizá tengan sentimientos, aunque estos sean ínfimos, ¿quién sabe?». Ahora bien, si jugamos esta baza, si decidimos que el robot hidra-cefálico es consciente... ¿por qué no lo puede ser España?

No pretendo afirmar que España tenga la misma organización funcional que el *homo sapiens*. Está claro que no. Pero existen muchas semejanzas. España tiene órganos de defensa. Tiene órganos de ingestión, se alimenta de petróleo y de cobre, etc. Excreta (polución) en vastas cantidades. ¿Acaso no es su organización funcional tan semejante a la de un mamífero como lo es la de una molesta mosca, a la que sí atribuimos dolor?

**PUTNAM, H.** (1988:**83-84**). *Razón, verdad e historia*. Madrid: Tecnos.

Por el momento no hemos puesto en duda lo que significa tener dos experiencias que comparten el mismo «carácter introspectivo». **En primer lugar, lo que nos parecen nuestras experiencias depende de una forma notoria de nuestras conceptualizaciones previas**, como muestra el que digamos que vemos una mesa circular aun cuando la observamos desde un ángulo.

Este último caso plantea la cuestión... de si tenemos «sense data elípticos» y *pensamos* que son circulares (salvo que estemos suficientemente entrenados en la práctica de la introspección) o tenemos «Gestalts» circulares y pensamientos que son elípticos debido a nuestro conocimiento de la teoría óptica.

... el carácter judicativo de la experiencia («veo la superficie circular de la mesa» corresponde igualmente a «registros» e «*inputs*» en el cerebro, pero si identificamos éstos con la sensación visual, el resultado puede ser otro relato según el cual no tenemos sense-data elípticos salvo que juzguemos que algo parece elíptico. ¿Por qué no afirmar que ambas versiones son legítimas? Como Goodman afirma con respecto al caso de un sujeto al que se le pide que describa el movimiento aparente.

**PUTNAM, H.** (1988:84). *Razón, verdad e historia*. Madrid: Tecnos

Lo que podemos hacer es especificar el tipo de términos, el vocabulario, que debe utilizar, pidiéndole que describa lo que ve, haciendo uso de términos perceptivo o fenoménicos más bien que términos físicos. Depare o no diferentes respuestas, este procedimiento arroja una luz completamente distinta sobre lo que sucede. **La necesidad de especificar el instrumento que debe utilizarse para configurar los hechos convierte en una insensatez cualquier tipo de identificación de lo físico con lo real y de lo perceptivo con la mera apariencia. Lo perceptivo es una versión bastante distorsionada de los hechos físicos en la misma medida en que lo físico es una versión sumamente artificial de los hechos perceptivos.**

**PUTNAM, H.** (1988:85). *Razón, verdad e historia*. Madrid: Tecnos.

Hemos examinado un conjunto de dificultades escépticas que no pretenden mostrar que la teoría de la identidad o la teoría de la correlación son erróneas ... sino que, si son verdaderas existe un gran número de sistemas alternativos para especificar todos los pormenores, de modo que nunca podemos saber cuál de ellos es verdadero. E ignorar esto significa ignorar la respuesta a muchos problemas tradicionales planteados por los escépticos, tales como si las rocas y otros objetos inanimados tienen *qualia*, *si los murciélagos y otras especies tienen o no el mismo tipo de qualia* que tenemos nosotros, si los grupos pueden sentir dolor, etc.

... Russell era realista con respecto a los *qualia* y con respecto a los universales. Además, consideraba que los *qualia* eran los universales paradigmáticos. **Un universal es, sobre todo, una forma en que las cosas pueden ser semejantes**; y opinaba que las semejanzas cualitativas entre las sensaciones de un mismo sujeto eran los ejemplos epistemológicamente más sencillos y fundamentales de «modos de semejanza entre las cosas». Los *qualia de Russell, son universales par excellence*.

**Un realista tradicional, sin embargo, concibe un universal como algo completamente bien-definido: las palabras pueden ser vagas, pero los universales no.** (Una palabra es vaga porque representa a un conjunto vago de conceptos, según dijo Gödel en cierta conversación, pero los conceptos están perfectamente definidos).

**PUTNAM, H.** (1988:85-86). *Razón, verdad e historia*. Madrid: Tecnos.

**De modo que, si los *qualia* son universales y los universales son por su naturaleza algo bien-definido, debe estar perfectamente-definido si cualquier cosa o evento** —incluyendo la mitad del cerebro escindido, una roca, una nación o un grupo, o cualquiera eventos que tengan un lugar en estos últimos— *manifiesta o no un *quale* determinado. Y si el *quale* se concibe como algo totalmente independiente del rol funcional que desempeña*, si se estima que es completamente contingente que el carácter cualitativo de la sensación de rojo sea el carácter cualitativo de algo que desempeña ese particular rol funcional, entonces parece ser una posibilidad lógica que el cerebro escindido o la roca tengan *quale*.

Un filósofo que comparta mi deseo de negar que todas estas y cada una de estas posibilidades tengan sentido... ha de dejar bien claro que no se adhiere a ninguna forma de behaviorismo. Decir que los *qualia* no son entidades bien definidas no es lo mismo que decir que no existen, que todo lo que existe es la conducta, o cualquier otra afirmación de esta índole. Muchas nociones son vagas y, con todo, poseen ciertos referentes claros. La noción de casa, por ejemplo, es una noción incorrectamente definida en el caso de los iglúes (¿consideramos que un iglú es una casa?)... Del mismo modo, el hecho de que no tengamos medios para ponderar y decidir si el lóbulo derecho es «consciente» no significa que no existan seres conscientes.

**PUTNAM, H.** (1988:86). *Razón, verdad e historia*. Madrid: Tecnos.

La semejanza cualitativa está definida hasta cierto punto: si experimento una sensación de rojo seguida por una sensación de verde, sé que he tenido sensaciones que no son semejantes (y lo sé sin necesidad de comparar sus roles funcionales), y he experimentado una sensación de rojo seguida por la «misma» sensación de rojo... que he experimentado sensaciones semejantes. Pero, para alguien que mantenga una perspectiva «internalista» con respecto a la verdad, no se sigue que tenga que cuestionarse en todos los casos si dos sensaciones (ni siquiera dos eventos arbitrarios) son cualitativamente semejantes o no.

Sea E el evento consistente en experimentar una sensación particular en un tiempo particular y E' algún evento físico en una roca. La suposición de que el carácter cualitativo de E (digamos rojo) podría ser idéntico o correlativo a una propiedad como (P1 o P2 o P3) (donde P3 es la propiedad de ser una roca) ofende a cualquier sensibilidad humana mínimamente sensata. La suposición de que E y E' pueden ser eventos «cualitativamente semejantes» es absurda. ... es absurda porque viola la máxima metodológica «no adscriba usted propiedades a un objeto sin tener alguna razón». Pues bien, aunque esta explicación fuese eficaz, estaría lejos de excluir la posibilidad de que las rocas tengan *qualia* ... Pero si esa transgresión es todo lo que tienen de erróneo las «hipótesis» que afirman que las rocas tienen *qualia*, entonces estamos en posición de afirmar: por lo que sabemos, es posible que las rocas tengan *qualia*, aunque *a priori* es sumamente improbable.

**PUTNAM, H.** (1988:86-87). *Razón, verdad e historia*. Madrid: Tecnos.

En realidad, la incoherencia de la hipótesis que afirma que las rocas tienen *qualia* es análoga a la incoherencia de la «hipótesis» de los cerebros en una cubeta. Como esta última, presupone una teoría mágica de la referencia. Cualquier ser humano mínimamente sensato consideraría que E y E' son tan poco semejantes que ni siquiera puede plantearse la cuestión de su «semejanza cualitativa» (en el sentido en que dos sensaciones pueden ser semejantes, esto es, afectándonos del mismo modo). Pero el realista metafísico, a pesar de que no niega esto último, piensa que E y E' podrían ser semejantes aun cuando sea un «disparate» pensar en esta posibilidad lógica. Y lo piensa porque se halla bajo la ilusión de que el hecho de experimentar la sensación de marras, con su carácter cualitativo, o con «su modo de afectar», con los pensamientos y juicios que la acompañan, ocasiona de algún modo que la expresión «el modo en que me afecta esta sensación» o «rojo», o «este quale») se refiera a un «universal» definido, esto es, a una propiedad absolutamente bien definida de acontecimientos metafísicos individuales. Pero las cosas no son así.

Aunque el caso es tan extraño que uno no puede estar seguro de nada, parece que incluso en esta caso (si nuestra interacción se da con el robot entero, y sólo raras veces con sus «neuronas» conscientes... podríamos comenzar a atribuirles consciencia; pero probablemente, nuestras opiniones están divididas. Si llegáramos a estar seguros de que los robots hidra-cefálicos son conscientes, ¿podríamos comenzar a estarlo, aunque siempre con ciertos escrúpulos, de que España lo es? No lo sé.

**PUTNAM, H.** (1988:87). *Razón, verdad e historia*. Madrid: Tecnos.

Con respecto a todos estos casos, recomiendo con énfasis la perspectiva que niega que en este punto exista algo oculto, algún hecho nouménico consistente en la mismidad real de las sensaciones o en la conciencia real de las entidades. Sólo existen hechos empíricos obvios: las rocas y las naciones son enormemente diferentes de las personas y de los animales, los distintos tipos de robots pertenecen a la clase de los objetos, etc. Las rocas y las naciones no son conscientes, y, con respecto a la noción de conciencia, esto es algo que consideramos como un hecho.

Esta concepción es tan perturbadora porque hace que nuestros estándares de aceptabilidad racional, justificación y, por último, de verdad, dependan de estándares de semejanza que son duda son producto de nuestra herencia cultural y biológica (por ejemplo, que hayamos interactuado o no con «robots inteligentes»). No obstante, ocurre algo parecido con respecto a buena parte del lenguaje que utilizamos en nuestra vida cotidiana, por ejemplo, con respecto a palabras como «persona», «nieve» y «marrón». Un realista que aceptase esta resolución de los enigmas acerca de los *qualia*, probablemente diría que «los *qualia* no existen realmente» o que los *qualia* pertenecen a nuestro «sistema conceptual de segundo orden». Pero **¿de qué sirve una noción de «existencia» que coloca a las cosas en el lado de lo *no-existente*?** Nuestro mundo es un mundo humano, *y la respuesta a qué cosas son conscientes o no, o a qué cosas experimentan sensaciones o no, o a qué cosas son cualitativamente semejantes o no, depende en última instancia de nuestros juicios humanos con respecto a la semejanza y la diferencia.*



**PUTNAM, H.** (1988:**88**). *Razón, verdad e historia*. Madrid: Tecnos.

## **5. DOS CONCEPCIONES DE LA RACIONALIDAD**

En los capítulos precedentes hemos estado hablando de racionalidad y de «aceptabilidad racional»...

El problema es análogo en otras áreas. Hace algunos años estudié el comportamiento de los términos de los géneros naturales, *oro*, por ejemplo, y llegue a la conclusión de que la extensión del término no se determina simplemente mediante una «batería de reglas semánticas» u otras normas institucionalizadas. Las normas pueden determinar que ciertos objetos son *ejemplos paradigmáticos* de oro, pero no determinan la plena extensión del término. Ni siquiera es imposible que un ejemplo paradigmático de oro resulte no ser realmente oro, y debería serlo si «ser de oro» estuviese definido simplemente por normas.

**Estamos dispuestos a considerar que algo pertenece a un determinado género, a pesar de que los *test* de que disponemos no basten para demostrarlo**, si ese algo resulta tener en todo caso la misma naturaleza que (o, de modo más vago, si es «**suficientemente semejante**» a) los ejemplos paradigmáticos (o que la gran mayoría de éstos). La naturaleza esencial, o la semejanza suficiente, depende tanto del género natural como del contexto (el té helado puede ser «agua» en un contexto, pero en otros no; más lo que cuenta para que algo sea oro es su composición última, puesto que (desde la antigua Grecia) se ha pensado que determinar el comportamiento legal de una sustancia consiste en determinar tal composición.

**PUTNAM, H.** (1988:**88-89**). *Razón, verdad e historia*. Madrid: Tecnos.

Es tentador adoptar la misma línea con respecto a la racionalidad, y afirmar que lo que determina que una creencia sea racional no son las normas de racionalidad de tal o cual cultura, sino la teoría ideal de la racionalidad, una teoría que nos ofrecería las condiciones necesarias y suficientes para que una creencia sea racional en las circunstancias relevantes y en cualquier mundo posible. Esta teoría tendría que dar cuenta de los ejemplos paradigmáticos de racionalidad, del mismo modo que una teoría ideal del oro lo hace con los ejemplos paradigmáticos de oro, pero podría ir más allá de éstos, proporcionándonos criterios que nos permitan entender casos que en la actualidad no podemos desentrañar por completo, tal y como nuestra actual teoría del oro nos permite entender casos que ni el más brillante de los antiguos griegos podría haber entendido. La propuesta de considerar «racional», «razonable», «justificado», etc., como términos de géneros naturales, tropieza con dificultad general de que las perspectivas de **encontrar generalizaciones poderosas con respecto a todas las ciencias racionalmente aceptables no son muy halagüeñas... pero, ¿tenemos alguna probabilidad de encontrar poderosas generalizaciones universales por toda instancia de «creencia racional justificada»?**

... vamos formulando gradualmente máximas (y no generalizaciones en excepción) basadas en juicios, que a menudo van acompañadas por razones o ejemplos ilustrativos, como por ejemplo «Sea usted amable con el extranjero, pues ya se sabe lo que significa ser extranjero en Egipto»... Estas máximas afectan y alteran a su vez nuestros juicios sobre casos individuales, de modo que pueden aparecer nuevas máximas que suplementen o modifiquen las anteriores.

**PUTNAM, H.** (1988:**90-91**). *Razón, verdad e historia*. Madrid: Tecnos.

Positivismo Lógico. Las positivistas no sólo confiaban en que los «lógicos de la ciencia» ... lograrían dejar constancia escrita de una descripción (pretendidamente exhaustiva) del método científico, sino que, según ellos, el «método científico» agotaba la propia racionalidad, y la contrastabilidad **mediante ese método agotaba la significatividad** («**el significado de una oración es su método de verificación**»); la lista o *canon* determinaría lo que tiene significación cognitiva. Los enunciados contrastables mediante métodos pertenecientes a la lista (los métodos de la matemática, la lógica y las ciencias empíricas) contarían como significados; los positivistas mantenían que todos los demás eran sólo «pseudo-enunciados», o sinsentidos disfrazados.

Wittgenstein sostenía que sin esas normas públicas, normas que son compartidas por un grupo y que constituyen una «forma de vida», no sería posible el lenguaje, ni tampoco el pensamiento. Según Wittgenstein, es absurdo preguntar si la verificación institucionalizada de la que he estado hablando desempeña un papel «realmente justificativo». En *Sobre la certeza*, Wittgenstein observa que los filósofos pueden proporcionarnos cientos de «justificaciones» epistemológicas del enunciado «Los gatos no brotan en los árboles»

**PUTNAM, H.** (1988:93-94). *Razón, verdad e historia*. Madrid: Tecnos.

... Pero ¿por qué era tan seductor para los filósofos proclamar que muchas cosas que son, en un sentido no-inteligible, «reglas del lenguaje», o consecuencias de las reglas del lenguaje, eran analíticas o «conceptualmente necesarias» o algo parecido? Creo que la respuesta es que los filósofos pensaban que la idea de que existe un conjunto definido de reglas del lenguaje y que éstas pueden establecer lo que es y lo que no es racional, tenía dos ventajas: (1) las «reglas del lenguaje» son prácticas constitutivas e institucionalizadas (o normas que subyacen a tales prácticas) y, como tales, tienen *status* «público» que he descrito. (2) Se afirmaba al mismo tiempo que sólo los filósofos (**y no los lingüistas**) podían descubrir esas cosas tan misteriosas. Fue una buena idea mientras duró, pero estaba condenada a ser refutada, y lo fue.

Llamaré **concepción criterial** a cualquier concepción de acuerdo con la cual la aceptabilidad racional se defina mediante normas institucionalizadas. Los positivistas lógicos, Wittgenstein ... y algunos filósofos del «lenguaje ordinario» de Oxford, aunque no todos, comparten una concepción criterial de la racionalidad...

El ámbito al que me referí al comienzo, el ámbito que refuta el principio de verificación lógico, es *incisivo* precisamente porque **refuta cualquier argumento a favor de una concepción criterial de la racionalidad, esto es, porque refuta a tesis de que nada es racionalmente aceptable si es criterialmente verificable.**

PUTNAM, H. (1988:94). *Razón, verdad e historia*. Madrid: Tecnos.

... **no existe ninguna posición filosófica de importancia que pueda ser verificada de la forma conclusiva y culturalmente reconocida** que he descrito. En resumen, si es cierto que sólo pueden ser racionalmente aceptables aquellos enunciados que pueden verificarse criterialmente, este mismo enunciado no puede ser verificado criterialmente y, por tanto, no es racionalmente aceptable. **Si es que existe tal cosa como la racionalidad** —y nos comprometemos a creer en alguna noción de racionalidad al tomar parte en las actividades de hablar y argumentar— entonces la actividad de *argumentar* en favor de una posición que la identidad a (o convierte en un subconjunto propio de) lo que las normas institucionalizadas determinan ya como instancias de racionalidad, es una actividad que se autorrefuta. Pues **esas normas no pueden garantizar por si solas la corrección, o la probable corrección, de ningún argumento de esta índole**.

No estoy negando la posibilidad de argumentación y justificación racional en filosofía; más bien me he visto obligado a reconocer algo que probablemente es evidente para los legos, aunque no para los filósofos, a saber: no podemos apelar a normas públicas para decidir qué enunciados filosóficos son racionalmente argumentables y justificables. La afirmación que aún hoy es frecuente oír, la afirmación de que la filosofía es «análisis conceptual», que los propios conceptos determinan qué argumentos filosóficos son correctos, es (cuando se combina con la doctrina de que los conceptos son normas o reglas que subyacen a las prácticas lingüísticas públicas) tan sólo una forma encubierta de afirmar que **toda justificación racional en filosofía es criterial**, y que la verdad filosófica es (si exceptuamos los «errores garrafales») *públicamente demostrable* en la misma medida en que lo es la verdad científica. Tal opinión me parece poco razonable a la luz de toda la historia del problema, incluyendo la historia reciente.

**PUTNAM, H.** (1988:95). *Razón, verdad e historia*. Madrid: Tecnos.

Recordemos que los positivistas concedían que el principio de verificación carecía de «significado cognitivo». Afirmaban que era una propuesta, y como tal, no era verdadera ni falsa. No obstante, si que argumentaban a favor de su propuesta y sus argumentos eran (y no tenían por qué ser) puntos de partida. Así que el problema permanecía irresoluto.

En resumidas cuentas, lo que hicieron los positivistas lógicos y Wittgenstein (y quizá también el último Quine) fue *producir filosofías que no dejaban lugar a la actividad racional de la filosofía*. Sus concepciones autorrefutan por ese motivo, pero también porque el pequeño gambito del que he estado hablando representa un argumento significativo, del tipo que los filósofos denominan «argumento trascendental»: argumentar con respecto a la naturaleza de la racionalidad (la tarea por excellence de los filósofos) **es una actividad que presupone una noción de justificación más amplia que la positivista, y en realidad, más amplia que la noción de racionalidad criterial institucionalizada.**

**PUTNAM, H.** (1988:95-97 ). *Razón, verdad e historia*. Madrid: Tecnos.

## **¿POR QUÉ ES INCONSISTENTE EL RELATIVISMO?**

Es un tópico entre los filósofos que el relativismo (total) es inconsistente. Después de todo, ¿no es *obviamente contradictorio mantener* un punto de vista mientras se mantiene al mismo tiempo que ningún punto de vista es correcto, ni está más justificado que cualquier otro? ... Si cualquier punto de vista es tan bueno como cualquier otro, ¿por qué no puede ser el punto de vista que afirma que el relativismo es falso un punto de vista tan bueno como cualquier otro?

... ¿Por qué un relativista inteligente habría de reconocer que todas las opciones son igualmente verdaderas (para él). El relativismo no puede impedir que usted ... afirme que su opinión no es verdadera para usted (o no tiene justificación, etc.): pero si conserva su presencia de ánimo, puede espetarle que lo que es verdad para usted dista de ser tan relevante (para él) como lo que es verdad para él. Después de todo, ¿hay algún concepto de algo que sea más relevante que el que uno tiene? ¿Es realmente inconsistente considerar *verdad, justificado, etc.*, como nociones *relativas*.

La respuesta afirmativa requiere un argumento más elaborado que el argumento unilineal (y, no obstante, muy acertado) aducido por Garfinkel [cuando ustedes van, ya yo he vuelto]. Lo importante es darse cuenta que si todo es relativo, lo relativo también es relativo. Mas esto requiere una explicación.

**PUTNAM, H.** (1988:103). *Razón, verdad e historia*. Madrid: Tecnos.

A Wittgenstein le interesaba atacar una forma de relativismo conocida como «solipsismo metodológico». Un «**solipsista metodológico**» es un **no-realista o «verificacionista»** que está de acuerdo en que la verdad debe concebirse emparentada de alguna forma con la aceptabilidad racional, pero que **sostiene que toda justificación viene dada en última instancia en términos de experiencias de las que cada uno tiene un conocimiento *privado***. De este modo, yo tengo mi conocimiento de cuáles de mis experiencias verificarían que la nieve es blanca y que Bob Nozik tiene su conocimiento de cuáles de sus experiencias verificarían que la nieve es blanca: cada enunciado tiene un sentido diferente para cada sujeto pensante.

El argumento de Wittgenstein me parece un argumento excelente contra el relativismo en general. Consiste en que, en última instancia, el relativista no puede dar ningún sentido a la distinción entre *estar en lo cierto y creer que se está en lo cierto*, y esto significa que, al fin y al cabo, no hay diferencia alguna entre *afirmar y pensar*, por una parte, y *producir ruidos o imágenes mentales* por otra. Pero esto significa que (según esta concepción) *yo no soy un ser que piensa*, sino un mero animal. Mantener este punto de vista nos compromete con una especie de **suicidio mental**.

... El relativista podría adoptar la idea de que la verdad es una *idealización* de la aceptabilidad racional. Podría mantener que X es-verdadero-para-mi si «X *tiene justificación para mí*» fuese verdadero siempre que lo examinase con suficiente cuidado, razonado lo suficiente, etc.



**PUTNAM, H.** (1988:95-**103-104**). *Razón, verdad e historia*. Madrid: Tecnos.

Un realista metafísico puede considerar que los enunciados que versan sobre lo que sería el caso sí... son por sí mismos absolutamente verdaderos o falsos, independientemente de que en alguna ocasión *esté justificado* el aceptarlos o rechazarlos. Si el relativista interpreta de este modo tan realista los enunciados que versan sobre lo que creería bajo tales-y-cuales condiciones, reconoce una clase de verdades absolutas y deja de ser relativista.

Un no-realista o un realista «interno» concibe los enunciados condicionales de justificación (como entendemos todos los demás enunciados). Esto no significa que el realista «interno» abandone la distinción entre verdad y justificación, sino que aprehendemos la verdad (*justificación idealizada*) y como aprehendemos cualquier otro concepto; por medio de una comprensión (en gran medida implícita) de los factores que hacen racionalmente aceptable decir que algo es verdadero. ¿Puede el relativista interpretar de esta forma no-realista, o realista «interna», los enunciados que se refieren a lo que creería en condiciones ideales?

Recordemos que en la posición no-realista, ... asume una noción *objetiva* de aceptabilidad racional. El no-realista rechaza la opinión de que la verdad es una correspondencia con «un mundo prefabricado». Y es lo que le convierte en un no-realista (metafísico). Pero rechazar la teoría de la «verdad-correspondencia» no es del todo lo mismo que considerar *subjetivas* tanto la verdad como la aceptabilidad. Nelson Goodman, quien concibe la verdad y la aceptabilidad racional como tipos de un predicado más general, «corrección», que se aplica tanto a las obras de arte como a los enunciados, ha expuesto el problema de un modo muy sucinto.

**PUTNAM, H.** (1988:95-**104**). *Razón, verdad e historia*. Madrid: Tecnos. Brevemente entonces, la verdad de los enunciados y la corrección de las descripciones, ejemplificaciones, expresiones —de diseño, de dibujo, de lenguaje, de ritmo— es principalmente una cuestión de **ajuste**: ajuste a modos y maneras de organización. Las diferencias entre ajustar una versión a un mundo. Ajustar un mundo a una versión y ajustar una versión a otra u otras versiones, se van desvaneciendo a medida que se reconoce el papel de las versiones en la elaboración de los mundos. Y se considera que el saber y el entendimiento se dan más allá de la adquisición de creencias verdaderas, hasta el descubrimiento y la ideación de ajustes de todo tipo.

Todo el *propósito* del relativismo, su propia característica definitoria, es, sin embargo, negar la existencia de cualquier noción inteligible de «ajuste» *objetivo*. Así que el relativista no puede entender la teoría de la verdad en términos de justificación *objetivas*.

El intento de usar *condicionales* para explicar la distinción entre estar en lo *cierto* y *creer* que se está en lo cierto fracasa porque el relativista no tiene una noción objetiva de corrección para estos condicionales, como no la tiene para ningún otro tipo de enunciados.

... El relativista debe acabar negando que los pensamientos se refieran a algo, en el sentido realista o en el no-realista. En resumen, el relativista so se da cuenta de que la existencia de algún tipo de «corrección» objetiva es una presuposición del propio pensamiento.

**PUTNAM, H.** (1988:108). *Razón, verdad e historia*. Madrid: Tecnos.

## **HECHO Y VALOR**

El asunto de los hechos y valores, entendido de un modo suficientemente amplio, nos concierne a todos. En esto se diferencia netamente de muchos problemas filosóficos. La mayor parte de las mujeres y de los hombres cultos no se sienten en la obligación de adoptar una opinión definitiva con respecto al «problema» de si hay verdaderamente un mundo real o si sólo parece haberlo. Desde el punto de vista del modo de vida mayoritario... pero de elección opcional. Pero el problema de los hechos y de los valores es de elección forzosa. Cualquier persona reflexiva ha de tener una opción real con respecto a éste (opinión que puede coincidir con su opción nocional). Y si la elección del problema de los hechos y valores es edundante forzosa,\* ocurre que una particular solución a este problema se ha arrogado el *status* de institución cultural, y es la siguiente: la dicotomía «enunciados de hecho o enunciados de valor» es absoluta.

Al llamar a la dicotomía «institución cultural» **pretendo sugerir que, por desgracia, la respuesta en cuestión seguirá siendo aceptada durante bastante tiempo, ignorando lo que los filósofos puedan decir sobre ella y sin tener en cuenta si es o no correcta.** ... la dicotomía hecho-valor carece de base racional... aun así ... La opinión según la cual no hay base objetiva para decidir si las cosas son buenas o malas, o mejor o peores, se ha convertido en algo *institucionalizado*.

... voy a rehabilitar un argumento un tanto desacreditado en el debate sobre los hechos y valores, a saber el que defiende que esa distinción es al menos irremediabilmente difusa, ya que los propios enunciados fácticos y los procedimientos de investigación empírica con los que contamos para decidir si algo es o no es un hecho presuponen valores.

\*En original inglés (p. 127: “totally disjoint realms”

**PUTNAM, H.** (1988:108-109). *Razón, verdad e historia*. Madrid: Tecnos.

El descrédito de este argumento se debe a una réplica obvia. ... apela a la presuposición de valores por parte de la ciencia constituye una concesión preventiva: los defensores de la dicotomía hecho-valor conceden que la ciencia presupone valores, por ejemplo, presupone que queremos la verdad, pero sostienen que esos valores no son éticos.

Como hemos visto, la verdad no es una noción simple. La idea conforme a la cual la verdad es una copia pasiva de lo que «realmente existe» (con independencia de la mente y del discurso) se ha derrumbado bajo las críticas de Kant, Wittgenstein y otros filósofos, bien continúa ejerciendo una profunda influencia sobre nuestra forma de pensar.

Algunos filósofos, **apelando al principio de equivalencia —esto es, el principio que reza: *decir que un enunciado es verdadero equivale a afirmar en enunciado*—** para negar que existan problemas filosóficos con respecto a la verdad. **Otros apela a la obra de Tarski**, el lógico que mostró de qué forma, dado un lenguaje formalizado (una notación formal en la que expresa ciertos enunciados, empleando la lógica simbólica), **podemos definir «verdadero» para ese lenguaje en un lenguaje de orden superior (denominado «metalenguaje»)**.

**PUTNAM, H.** (1988:109). *Razón, verdad e historia*. Madrid: Tecnos.

La propia obra de Tarski está basada en el **principio de equivalencia**: de hecho, su criterio para obtener una definición adecuada de «verdad» **consistía en que ésta había de ofrecer como teoremas del meta-lenguaje todas aquellas oraciones de la forma «P» es verdadera si y sólo si P (donde P es una oración de la notación formal en cuestión), por ejemplo,**

(T) «La nieve es blanca» es verdadera si y sólo si la nieve es blanca

No obstante, el principio de equivalencia es filosóficamente neutral, como lo es la obra de Tarski. Según *cualquier* teoría de la verdad, «La nieve es blanca» es equivalente a «La nieve es blanca es verdadera.»

Los filósofos positivistas replicarán que si usted sabe (T), sabe lo que significa "La nieve es blanca" significa La nieve es blanca. Y añadirán que si usted no comprendiese «nieve» y «blanca», se hallaría en un serio apuro. Pero el problema no es que no comprendemos «La nieve es blanca»; el problema es que *no comprendemos qué comprender*. ...

**PUTNAM, H.** (1988:110-113). *Razón, verdad e historia*. Madrid: Tecnos.

Si la idea de comparar nuestro sistema de creencias con una realidad conceptualizada para ver si se «emparejan» no tiene sentido, entonces la afirmación: «la ciencia busca descubrir la verdad» solo puede significar que la ciencia constituye una imagen del mundo que, en el límite ideal, satisface ciertos criterios de aceptabilidad racional. Que la ciencia busca construir una imagen del mundo que sea verdadera es en si mismo un enunciado verdadero, un enunciado verdadero, formal y casi vacío; sólo los criterios de aceptabilidad racional implícitos en la ciencia dotan de contenido material a sus objetivos. En resumen, estoy diciendo que la respuesta a la posición del oponente imaginario, **aquella que defiende que el único objetivo de la ciencia es descubrir la verdad... es que la verdad no es la cuestión de fondo: la verdad misma obtiene su vida a partir de nuestros criterios de aceptabilidad racional, y debemos examinar éstos si deseamos descubrir os valores que están efectivamente implícitos en la ciencia.**

Nuestra actividad científica es construir una representación del mundo que tenga las siguientes características: **eficacia instrumental, coherencia, alcance comprensivo y simplicidad funcional. Pero ¿por qué razón?**

**PUTNAM, H.** (1988:110-113). *Razón, verdad e historia*. Madrid: Tecnos.

Desde luego, si el realismo metafísico fuera correcto, si fuese posible contemplar el objetivo de la ciencia como intento de «emparejar» nuestro mundo nocional con el mundo en-si-mismo, entonces se podría afirmar que sólo nos interesa la coherencia, el alcance comprensivo, la simplicidad funcional y la eficacia instrumental en tanto que instrumentos que sirven al fin de lograr este «emparejamiento». Pero en concepto de «emparejamiento trascendental» entre nuestras presuposiciones y el mundo en si mismo no tiene sentido. Con todo, negar que busquemos este tipo de igualación metafísica con un mundo metafísico no es negar que busquemos este tipo usual de ajuste empírico (decidida por nuestros criterios de aceptabilidad racional) con un mundo empírico. Pero el mundo empírico, en oposición con al mundo nouménico, depende de nosotros criterios de aceptabilidad racional para elaborar una imagen teórica del «»mundo empírico y conforme se desarrolla esta imagen revisamos bajo su luz nuestros propios criterio de aceptabilidad racional, y así sucesiva e ininterrumpidamente. En mis otros libros he puesto énfasis en la dependencia de nuestros métodos con respecto a nuestra imagen del mundo. **Lo que quiero resaltar aquí** es la otra cara de la moneda, **la dependencia del mundo empírico con respecto a nuestros criterios de aceptabilidad racional**. Lo que trato de **afirmar es que hasta para tener mundo empírico debemos tener criterios de racionalidad**, y que éstos revelan parte de nuestra concepción de una inteligencia especulativa óptima. **En resumen, estoy afirmando que «el mundo real» depende de nuestros valores (y, una vez más, también al contrario).**

**PUTNAM, H.** (1988:114-115). *Razón, verdad e historia*. Madrid: Tecnos.

El hecho de que la ciencia no sea, como se ha pensado, «neutral en cuanto a los valores» no muestra ni que los valores «éticos» sean objetivos, ni que la ética pueda ser una ciencia, claro está. De hecho, no existe ninguna probabilidad de alcanzar una «ciencia» de la ética, sea en el sentido de una ciencia de laboratorio o en el sentido de una ciencia deductiva.

... Por el hecho de que la aceptabilidad racional en las ciencias exactas... dependen de *virtudes cognitivas* como la «coherencia» y la «simplicidad funcional» muestra que al menos algunos términos de valor representan propiedades de las cosas a las que aplican, y no precisamente sentimientos de la persona que los usa.

Si los términos «coherente» y «simple» no están por *propiedades* de las teorías, ni siquiera por propiedades difusa o imperfectamente definidas, sino sólo por «actitudes» que ciertas personas mantienen hacia esas teorías, entonces los términos que guardan relación con la aceptabilidad racional, tales como «justificado», «bien confirmado» o «la mejor explicación de que disponemos», deben ser subjetivos casi por completo: pues la aceptabilidad racional no puede ser más objetiva que los parámetros de los que depende. No obstante... la concepción que mantiene que la aceptabilidad racional es sencillamente subjetiva se refuta a sí misma. De modo que nos vemos obligados a concluir que al menos estos términos de valor tienen algún tipo de aplicación objetiva, esto es, algún tipo de condiciones objetivas de justificación.

... **No hay ninguna concepción neutral de la racionalidad a la que apelar.**



**PUTNAM, H.** (1988:116). *Razón, verdad e historia*. Madrid: Tecnos.

## **LA RACIONALIDAD EN OTRAS ÁREAS**

Hasta ahora sólo hemos examinado los estándares de la **aceptabilidad racional** en un sentido literal; estándares que nos dicen cuando hemos de aceptar enunciados y cuando no hemos de aceptarlos. Ahora bien, en un sentido más amplio, los estándares de aceptabilidad tienen que ver con el modo en que juzgamos no sólo la verdad y la falsedad de los sistemas de enunciados, sino también su *adecuación y perspicuidad*. Hay otros motivos —estrictamente cognitivos— por los que un sistema de enunciados puede ofrecernos una descripción que no resulte suficientemente satisfactoria, además de por ser falso.

De haberlo querido, podría haber puesto de relieve este punto en conexión con la ciencia teórica. Podría haber señalado que el interés de la ciencia exacta no es sólo descubrir enunciados verdaderos, incluso enunciados verdaderos y universales («Leyes»), sino encontrar enunciados que sean **verdaderos y relevantes**. Y el concepto de relevancia arrastra tras de sí un amplio grupo de intereses y valores. Con todo, ello sólo habría sido apuntar que nuestro conocimiento del mundo presupone valores, y no establecer la afirmación más radical de que lo que cuenta como mundo real depende de nuestros valores.

**PUTNAM, H.** (1988:116-117). *Razón, verdad e historia*. Madrid: Tecnos.

Cuando llegamos a la racionalidad perceptiva, es esto es, a las técnicas y criterios implícitos sobre cuya base decidimos si alguien puede ofrecernos un informe verdadero, adecuado y perspicaz hasta de los **hechos perceptivos** más simples, vemos que entran en juego un gran número de factores. **Recientemente los psicólogos han hecho hincapié en que hasta los casos más simples de percepción envuelven bastante construcción teórica.** Esto no sólo es cierto a nivel neurofisiológico, sino también a nivel cultural. Una persona perteneciente a una cultura que carece de muebles podría ofrecer cierto tipo de descripción de una habitación, pero si no sabe que es una mesa, una silla o un pupitre, su descripción difícilmente comunicaría la información sobre la habitación que un miembro de nuestra cultura esperaría obtener. **Su descripción podría constar de enunciados verdaderos, y aún así, no ser la descripción adecuada. ...**

Lo que es cierto en un nivel tan simple... lo es también en el nivel de la descripción de las relaciones y situaciones interpersonales. Examinemos los términos que usamos todos los días para describir o que nos parecen las demás personas, por ejemplo, *amables o poco amables*. Estos términos pueden usarse, cómo no, para elogiar o censurar; y una de las muchas distinciones que se han confundido bajo el rótulo general «distinción hecho-valor» es la distinción entre usar una expresión lingüística para describir y usarla para elogiar o censurar. Mas esta distinción no puede trazarse sobre la base del vocabulario. El juicio que afirma que alguien es poco amable puede emplearse para censurar; pero puede emplearse sencillamente para describir, y también para explicar o predecir.

**PUTNAM, H.** (1988:117). *Razón, verdad e historia*. Madrid: Tecnos.

Por ejemplo, podría advertirle al lector: «No deje que Hernández hiera sus sentimientos. Su forma de hablar puede insinuar cierta antipatía, pero esta impresión errónea es bastante normal. Es probable que se siga comportando de un modo que hiera sus sentimientos, sienta lo que siente por usted. Es un hombre poco amable, pero no piense que tiene algo contra usted».

En este imaginario discurso, estoy empleando las palabras «poco amable» sin el propósito de censurar a Hernández, sino con la intención de predecir y explicar a alguien su comportamiento. Y tanto la predicción como la explicación pueden ser perfectamente correctas. De forma similar «celoso» puede ser un término de censura, pero también puede usarse sin intención de censurar...

El uso de la palabra «amable» me parece un buen ejemplo de cómo la distinción hecho-valor es desesperadamente difusa en el mundo y en el lenguaje reales. ... Cuando concebimos hechos y valores como cosas independientes, lo típico es pensar que los «hechos» se enuncian de alguna jerga fiscalista o burocrática, y que los «valores» necesitan de términos de valor más abstractos, por ejemplo, «bueno», «malo». Pero es difícil seguir manteniendo la independencia de los hechos respecto a los valores cuando los mismos hechos son del tipo «poco amable»... Así como criticamos a un descriptor que no emplee los conceptos de mesa y silla cuando se le exige su uso, quien no observe que alguien es *amable* o espontáneo puede exponerse a ser tachado de imperceptivo o superficial; su descripción no es una descripción adecuada.

**PUTNAM, H. (1988:117-118).** *Razón, verdad e historia.* Madrid: Tecnos.

## **LOS SUPERBENTHAMITAS**

... En resumen, no es cierto que «coincidamos con los superbenthamitas en los hechos y discrepamos con respecto a los valores». En relación con casi todas las situaciones interpersonales, nuestra descripción de los hechos será completamente distinta de la suya. Aun cuando ninguno de sus enunciados acerca de la situación fuera *falso*, no consideraríamos a su descripción ni adecuada ni perspicua. Y ellos imputarían los mismos defectos a la nuestra. **En resumen, aun cuando dejemos a un lado nuestro «desacuerdo con respecto a los valores» no podríamos concebir su representación del mundo como una representación completa y racionalmente aceptable.**

## **SUBJETIVISMO CON RESPECTO A LA BONDAD**

... La necesidad de «premisas morales» para extraer conclusiones morales a partir de enunciados «fácticos» muestra que los enunciados éticos no son racionalmente justificable. **Esta imagen de la ética como una especie de pirámide invertida, con un vértice (que carece de apoyos) consistente en «axiomas éticos» y que cimenta todo nuestro cuerpo de creencias y pensamientos morales, es ingenua.** Nadie ha logrado imponer una estructura axiomática sobre la ética... En realidad, la ética, la matemática y el discurso sobre objetos materiales presuponen conceptos, no «axiomas». **Los conceptos se usan en la observación y en las generalizaciones, y están legitimados por el éxito que logramos al utilizarlos en esas actividades.**

**PUTNAM, H.** (1988:119-120). *Razón, verdad e historia*. Madrid: Tecnos.

Un ataque más sofisticado a la idea de objetividad ética admite que nuestras creencias éticas descansan en observaciones de casos específicos, «intuiciones», máximas generales, etc., y no en un cúmulo de «axiomas éticos» arbitrarios, pero formula la acusación que establece que las propias observaciones éticas están afectadas por una incurable dolencia: la **proyección**.

De acuerdo con esta descripción, los humanos somos por naturaleza compasivos, si bien no siempre... nos horrorizamos (a veces). Por el mecanismo psicológico de la «proyección» nos lleva a experimentar la cualidad sentida como una cualidad de la propia acción: decimos «Fue un acto horroroso», cuando realmente deberíamos decir: «Mi reacción fue sentirme horrorizado». De esta manera vamos construyendo un cuerpo de lo que nosotros consideramos observaciones éticas, pero que son en realidad tan sólo observaciones con respecto a nuestros propios *sentimientos éticos subjetivos*.

... Hume postuló una tendencia humana a lo que llamó «simpatía», tendencia que ha ido ampliándose gradualmente bajo la influencia de la cultura. Los sociobiólogos contemporáneos postulan un instinto al que llaman «altruismo» y hablan de «genes altruistas»: Pero la idea clave sigue siendo la misma: **hay sentimientos éticos, pero no propiedades objetivas de valor**.

**PUTNAM, H.** (1988:120-121). *Razón, verdad e historia*. Madrid: Tecnos.

Ya hemos visto que esto no es cierto: hay al menos algunos valores objetivos, por ejemplo, la *justificación*. Con todo, podría afirmarse que los valores *éticos* son subjetivos, mientras que los valores *cognitivos* son objetivos; pero el argumento que establece que es absolutamente imposible que hayan valores objetivos ha sido refutado.

... Esto parece extraño, ¿no es el subjetivismo lo contrario del realismo metafísico? Si alguien cree que sí es así, también creerá que cualquier argumento contra el realismo metafísico debe respaldar al subjetivismo; la estrategia que voy a seguir, a saber, emplear el mismo argumento tanto contra el realismo metafísico como el subjetivismo, le parecerá una estrategia imposible.

Pero, de hecho, el realismo metafísico y el subjetivismo no son simples «contrarios». En nuestros días tendemos a ser demasiado realistas con respecto a la física y demasiado subjetivistas con respecto a la ética, y estas tendencias están relacionadas. Es porque somos demasiado realistas en física, porque consideramos a esta última (o a alguna hipotética física futura) como la Teoría Verdadera, y no simplemente como una descripción racionalmente aceptable, adecuada a ciertos problemas y propósitos, por lo que tendemos a ser subjetivistas con respecto a las descripciones que no podemos «reducir» a la física. Así mismo, llegar a ser menos realista con respecto a la física y menos subjetivista con respecto a la ética son tendencias que también guardan relación.

**PUTNAM, H.** (1988:122-123). *Razón, verdad e historia*. Madrid: Tecnos.

Kurt Gödel creía que «la intuición matemática» era análoga a la **percepción**. Los objetos matemáticos (a los que denominó «conceptos») están *ahí fuera*, y nuestra intuición nos permite percibir intelectualmente esas entidades platónicas; no obstante, pocos matemáticos se comprometerían con esta metafísica. La comparación gödeliana de la intuición matemática con la percepción revela una idea demasiado simple de esta última.. La visión no nos ofrece un acceso directo a un mundo prefabricado, sino que nos ofrece una descripción de los objetos según la cual estos últimos están en parte estructurados y constituidos por la propia visión. Si consideramos que el arco iris del físico es el arco iris «en sí mismo», entonces el arco iris «en sí mismo» no tiene franjas (el análisis espectroscópico arroja una distribución uniforme de frecuencias); las franjas roja, anaranjada, amarilla, verde, azul y violeta **son características del arco iris perceptivo, no del arco iris físico**. EL arco iris perceptivo depende de la naturaleza de nuestro propio aparato perceptivo —de **nuestra «construcción visual del mundo»**, según la denominación de Nelson Goodman. (Los «objetos» del mundo físico también dependen de nuestra construcción-del-mundo, como muestra la plétora de versiones radicalmente diferentes que los físicos construyen a partir de los «mismos» objetos). A pesar de todo, no pensemos que la visión sea *defectuosa* porque se vean franjas en el arco iris; quien no pudiese verlas sí tendría una visión defectuosa. **La visión está avalada por su capacidad para facilitarnos una descripción que se ajusta a los objetos para nosotros, y no las cosas metafísicamente en-sí-mismas. La visión es buena cuando nos permite ver el mundo tal como es», esto es, un mundo humano y funcional, creado en parte por la propia visión.**

**PUTNAM, H.** (1988:123). *Razón, verdad e historia*. Madrid: Tecnos.

... En resumen, la intuición matemática es buena cuando nos permite ver los hechos matemáticos «tal y como son», esto es, como son en un mundo matemático contraído por la práctica matemática humana, incluyendo es esta práctica la aplicación de la matemática a otras materias. d

Una descripción psicológica o fisiológica de la visión no puede decirnos si la visión de las franjas en el arco iris cuenta como una «visión correcta» o no. Tampoco una descripción psicológica o fisiológica del proceso que acontece en el cerebro cuando «aprehende» el principio de inducción matemática, podría (aún menos) decirnos si el principio es verdadero o no. Una vez se admite este punto, no debería sorprendernos que una descripción del proceso cerebral que acontece cuando uno «ve» una acción injusta no podemos decirnos si la acción es realmente justa.

El discurso sobre la «percepción» moral, como el discurso sobre la intuición matemática, o el discurso sobre la referencia y la comprensión, no es reducible a una imagen fiscalista del mundo. Esto no significa que la física sea «incompleta» ... **ninguna teoría o descripción puede ser completa para todos los propósitos**. Si la irreductibilidad de la ética a la física muestra que los valores son proyecciones, entonces los colores también lo son. Y también los números naturales. Y, puestas así las cosas, también «el mundo físico». Por ser una proyección, en este sentido, no es lo mismo que ser subjetivo.



**PUTNAM, H.** (1988:124). *Razón, verdad e historia*. Madrid: Tecnos.

## **AUTORITARISMO Y PLURALISMO**

He estado argumentando que es necesario tener estándares de aceptación racional hasta para tener un mundo, sea un mundo de «hechos empíricos» o un mundo de «hechos de valor» (un mundo en el que hay belleza y tragedia). No hará falta decir que no es posible tener criterios de racionalidad sin aceptarlos, o manteniéndose a distancia de ellos. (La clase de escepticismo consistente en negar que uno tiene algún estándar de racionalidad implica admitir que uno carece de conceptos. Como reconoció Sexto Empírico, este tipo de empirismo es, en última instancia, inexpresable en lenguaje). Tenemos tanto derecho a considerar que algunas inclinaciones «evaluativas» son enfermizas (y todos lo hacemos) como a considerar que lo son algunas inclinaciones «cognitivas». Pero esto no significa rechazar el pluralismo, ni tampoco comprometerse con el autoritarismo.

... (Bernard Williams ha puesto de manifiesto la especial importancia de una clase de casos indeterminados: los casos en los que todas las alternativas son tan atroces que ninguna de ellas sería elegida por una persona idealmente racional y sabia). **Y no hace falta decir que en la ética existen muchos aspectos que son relativos al contexto.**

**PUTNAM, H.** (1988:125). *Razón, verdad e historia*. Madrid: Tecnos.

Tampoco debe confundirse el compromiso con objetividad ética y lo que es una cuestión muy diferente, el compromiso con el autoritarismo ético o moral. Quizá sea esta la confusión la que ha llevado a que un destacado filósofo se adhiera a lo que él mismo considera como una versión limitada del «no-cognitivismo», y a afirmar que «Con respecto a lo que cada hombre es “vivir plenamente” la autoridad final debe ser el propio hombre». (Notemos la ambigüedad que hay en la expresión «la autoridad final»: quiere decir la autoridad política final? ¿Quiere decir autoridad epistémica final? ¿O quiere decir que no existen medios para ponderar y decidir, tal como sugiere su uso del término «no-cognitivismo»?). El respeto a las personas como agentes morales autónomos requiere que les reconozcamos el derecho a elegir su punto de vista moral, por muy repulsiva que encontremos su elección. De acuerdo con la filosofía del liberalismo político, también requiere que insistamos en que el gobierno no se asegure de antemano las elecciones morales individuales instaurando una religión o una moralidad de Estado. La oposición incondicional hacia toda forma de autoritarismo político y moral no debería comprometernos con el relativismo o con el escepticismo moral. La razón por la que el gobierno hace mal en dictar una moralidad para el ciudadano no es que no existan medios de ponderar qué formas de vida son satisfactorias y qué formas de vida no lo son. O son moralmente inicuas por algún otro motivo. (Si no hubiese algo como la incorrección moral, el gobierno no haría mal en imponer las elecciones morales). El hecho de que mucha gente tema que si conceden en voz alta algún tipo de objetividad moral encontrarán un gobierno que les haga tragar a la fuerza su noción de objetividad moral, constituye sin duda una de las razones por las que tantas personas subscriben un subjetivismo moral que realmente no aprueban.

**PUTNAM, H.** (1988:126). *Razón, verdad e historia*. Madrid: Tecnos.

## **7. RAZÓN E HISTORIA**

...algunos han seguido la moda cultural, sosteniendo que en realidad no hay conocimiento fuera de las ciencias exactas (y de las ciencias sociales en la medida en que imitan con éxito a las ciencias exactas, y sólo en esa medida). Esta opinión puede cobrar la forma del positivismo o del materialismo, o de alguna combinación de ambos. Otros han intentado argumentar que la ciencia también es «subjetiva» y arbitraria ... Otros, —por ejemplo, los filósofos marxistas y los filósofos de la religión— adoptan una especie de doble contabilidad, dejando los problemas técnicos a las ciencias exactas y a la ingeniería, y las cuestiones éticas e ideológicas a un tribunal diferente: el partido, la utopía futura, la iglesia... Y si nos enfrentamos a un auténtico problema en este ámbito es precisamente porque nos sentimos incómodos.

El problema es en cierto modo irreal, claro: la misma persona que sostiene que las opiniones éticas y políticas son inverificables, argumenta apasionadamente en favor de sus opiniones éticas y políticas.

**PUTNAM, H.** (1988:145-146). *Razón, verdad e historia*. Madrid: Tecnos.

## **8. EL IMPACTO DE LA CIENCIA EN LAS CONCEPCIONES MODERNAS DE LA RACIONALIDAD**

Si la discusión que hemos examinado... parece inconclusiva quizá sea porque siempre presupone cierto tipo de prioridad de la racionalidad sobre la bondad. La pregunta es siempre si hay algún sentido en el que pueda llamarse «irracionalidad» elegir un mal fin, como si la bondad estuviera sometida a juicio y la racionalidad fuera el juez. Adoptar esta postura, especialmente cuando nuestros presupuestos con respecto a la racionalidad son, en gran parte, una colección inexaminada de mitos y prejuicios culturales, es juzgar de antemano la cuestión del *status* de los juicios de valor. Propongo invertir los términos de la comparación en lo que resta de este ensayo, no preguntando ¿cuán racional es la bondad?, sino ¿por qué es bueno ser racional? La pregunta por el valor de la racionalidad nos obligará a clarificar su naturaleza y a arrojar luz sobre los presupuestos que tendemos a admitir con respecto a la racionalidad, y puede permitirnos comprender cuál es la causa de nuestros errores al afrontar la pregunta anterior.

Recordemos que cuando Max Weber introdujo la moderna distinción entre hechos y valores, su argumento en contra de la objetividad de los juicios de valor consistía precisamente en que no es posible establecer la verdad de un juicio de valor de un modo satisfactorio para toda *posible persona racional*. Ya desde el principio era la imposibilidad, o la pretendida imposibilidad de prueba racional, la que arrojaba una luz un tanto sospechosa sobre los juicios de valor. **La racionalidad ha estado llevando a juicio al valor desde hace mucho tiempo.** Y en este contexto racionalidad siempre significa racionalidad científica, se afirma que los resultados de la ciencia positiva son los únicos que pueden establecerse satisfactoriamente para todas las personas racionales.

**PUTNAM, H.** (1988:147-148). *Razón, verdad e historia*. Madrid: Tecnos.

La formulación de Weber sugiere que, de algún modo, primero hacemos un recuento de aquellos miembros de la población que son racionales y, después, observamos si todos ellos están dispuestos a admitir que algún juicio de valor es verdadero o no lo es. Pero, en realidad, las cosas no son así. Todo lo que muestran los ejemplos de Weber ... es que los juicios de valor no pueden ser verificados con el beneplácito de todas las personas cultas o inteligentes (que de ninguna manera son las mismas que todas las personas racionales). El argumento de Weber es, de forma velada, un argumento mayoritarista: apela al hecho de que podemos obtener el acuerdo de la gente culta en relación con la «ciencia positiva», mientras no podemos obtener acuerdo con respecto a los valores éticos.

El instrumentalismo no constituye por sí mismo una concepción de la racionalidad intelectualmente sostenible, desde luego. Nadie duda que los resultados científicos tienen un enorme valor práctico, pero, como ya sabemos, ninguna persona culta sostiene que el único valor de la ciencia lo constituyan sus aplicaciones. Y aun cuando se valorase únicamente por sus aplicaciones, ¿por qué ha de valorarse la *racionalidad* por sus aplicaciones?.. también es valioso saber qué fines hemos de escoger. No es sorprendente que la verdad de los juicios de valor no pueda ser «racionalmente demostrada» si la «verificación racional» está limitada, por definición, al establecimiento de las conexiones medios-fines. ... ¿por qué hemos de mantener una concepción de la racionalidad tan estrecha?

**PUTNAM, H.** (1988:148-149). *Razón, verdad e historia*. Madrid: Tecnos.

El mayoritarismo también es insostenible. Por supuesto que es agradable obtener el acuerdo con respecto a los que uno considera que es verdadero. ... Más la gente ha vivido durante siglos con el incómodo conocimiento de que en ciertos asuntos uno tiene que confiar en su propio juicio, aun cuando difiera del juicio de la mayoría. ... No soy sociólogo y no intentaré investigar la cuestión de por qué la sociedad industrial, tanto en sus versiones capitalistas como socialistas, ha de estar tan ligada a los temas del éxito instrumental y del consentimiento mayoritario...

Si bien la concepción de racionalidad que acabamos de describir, a saber, la idea de que la racionalidad consiste en modelos (cuya naturaleza queda en el aire) que dan por resultado el descubrimiento de conexiones efectivas entre medios y fines en el establecimiento «público» de estas conexiones, tal como está, es intelectualmente insostenible, no han faltado intentos filosóficos por hacerla respetable. Uno de esos intentos nace del más antiguo empirismo de Locke, Berkeley y Hume. Por la época de Mili, este empirismo se había solidificado en lo que los filósofos llaman **fenomenalismo: la doctrina que afirma que, en realidad, sólo podemos referirnos a sensaciones**. Según esta concepción, hasta los objetos cotidianos, mesas y sillas, son de hecho solamente conjuntos de regularidades objetivas de sensaciones humanas actuales posibles. Como expuso Mili, los objetos físicos son «posibilidades permanentes de sensación». **Otro modo de exponer la misma idea es decir que todo discurso que parece referirse al mundo físico es únicamente un discurso referente a sensaciones, si bien muy derivado.**

**PUTNAM, H.** (1988:149-150). *Razón, verdad e historia*. Madrid: Tecnos.

Este punto de vista tenía una virtud, a los ojos de sus adalides, de que les permitía decir con claridad cual era el contenido no sólo de la ciencia, sino también de cualquier discurso cognitivamente significativo. Cualquier teoría científica es sólo un medio «económico» de establecer una cantidad de hechos de la forma: *si realiza tales y tales acciones, entonces obtendrá tales y tales experiencias*. El partidario de esta concepción esta concepción no tiene por qué defender la insostenible afirmación de que los científicos sólo están interesados en las aplicaciones, o en la obtención de metas prácticas, y no en el conocimiento por sí mismo. El fenomenalismo no tiene por qué negar nuestro deseo de conocer la naturaleza de los agujeros negros, de saber si hubo o no Big-Bang, o de saber cuál fue el auténtico origen del *homo sapiens*. Queremos saber todas estas cosas, y no sólo porque su conocimiento nos permita construir mejores máquinas. Por conocer estas cosas es... conocer una gran cantidad de hechos de la forma: *si realiza tales y tales acciones, entonces tendrá tales y tales experiencias*. Cualquiera que sea el motivo por el que nos interesen todos los hechos son *en última instancia* instrumentales. Al mismo tiempo, no parece que haya manera de descifrar la afirmación de que llamar «bueno» es llevar a cabo una predicción de la forma: *si realiza tales y tales acciones, entonces tendrás tales y tales experiencias*. De modo que, según esta concepción, los enunciados con respecto a la bondad o a la maldad de algo no tiene significado cognitivo; dicho con palabras de los **empiristas lógicos del siglo veinte, tales enunciados son «puramente emotivos»**. . No obstante **el fenomenalismo fracasó en dos puntos...**

**PUTNAM, H.** (1988:149-150). *Razón, verdad e historia*. Madrid: Tecnos.

Las teorías científicas, entendidas como todos integrados, nos inducen a esperar que tendremos ciertas experiencias si llevamos a cabo ciertas acciones... ya ha sido abandonada como inaceptable tipo de reduccionismo. **En segundo lugar, las sensaciones son necesariamente objetos privados...** «¿cómo sabe usted que tal persona asocia a sus descripciones las mismas sensaciones que usted?», nos encontramos en medio de un problema epistemológico. Si el contenido de la ciencia consiste en predicciones con respecto a qué sensaciones tendrá un ser racional si realiza tales acciones, entonces para saber lo que ese mismo contenido significa, si nos encontramos ante un extraterrestre, habríamos de ser capaces de decir si tiene o no las mismas sensaciones que nosotros, etc. Por esta razón filósofos como Rudolph Carnap y Sir Karl Popper insistieron en que las predicciones observacionales de la ciencia deben establecerse de esta forma: si alguien realiza *tales y tales acciones, entonces tendrán lugar tales y tales acontecimientos públicamente observables, donde tanto las acciones como los acontecimientos observables esperados deben describirse en términos de objetos «públicos», por ejemplo, lecturas de contador, y no en términos de objetos privados tales como las sensaciones.*

En suma, el antiguo empirismo, o el fenomenalismo, parecía proporcionarnos un pulcro criterio de significación cognitiva: **un enunciado es cognitivamente significativo si es traducible a un enunciado acerca de sensaciones.** Pero resulta que o bien la noción de «traducción» es desesperadamente vaga, o bien los propios enunciados de la ciencia dejan de satisfacer ese criterio de significación cognitiva.



**PUTNAM, H.** (1988:150-151). *Razón, verdad e historia*. Madrid: Tecnos.

No obstante, nuestro propósito original fue considerar qué respuestas se habían dado a la pregunta «¿**Por qué es bueno ser racional?**» La primera respuesta que examinamos y rechazamos por ser demasiado estrecha afirmaba que la racionalidad nos permite descubrir conexiones fiables entre medios y fines. El fenomenalismo entró en escena porque, de ser verdadero, disolvería el conflicto entre interés en una teoría científica por sus consecuencias instrumentales y el interés por aprender lo que esta teoría nos cuenta con respecto a los procesos culturales. ... Todo conocimiento que merezca este nombre se habría mostrado como un conocimiento de conexiones de medios-fines. Sólo llamamos «prácticos» a nuestros intereses cuando estamos interesados en la conexión medios-fines, porque esperamos explotarla de cara a la obtención de alguna meta, y los llamamos «teóricos» cuando nos interesa conocer la conexión medios-fines por pura curiosidad. Como hemos señalado, este intento de reducir cada enunciado científico a un enunciado de la forma *si realiza la acción A entonces obtendrá el resultado B* ha fracasado.

**PUTNAM, H.** (1988:152). *Razón, verdad e historia*. Madrid: Tecnos.

La aceptación de este dogma [observacional] epistemológico hace que el que las creencias racionales sean susceptibles de comprobación *pública* sea parte de la definición de racionalidad. Y ello es conveniente, pues, hace innecesario proporcionar un argumento en favor de esta propuesta. Quizá el argumento consista en el fondo afirmar que lo que no es públicamente comprobable puede convertirse en materia de desacuerdo, y por otra parte, que siempre que hay un desacuerdo irresoluble no hay corrección o incorrección. Pero esto sería asumir lo que he llamado mayoritarismo, es decir, la idea de que en la propia noción de racionalidad está implícito que lo que es racionalmente verificable es verificable con el beneplácito de la aplastante mayoría.

**PUTNAM, H.** (1988:165). *Razón, verdad e historia*. Madrid: Tecnos.

## 9. **HECHOS, VALORES Y COGNICIÓN**

... Así, pues, «**racionalmente aceptable**» son **nociones interdependientes**, Y argumentaba que **ser racional implica tener un criterio de relevancia además de un criterio de racionalidad, y que en nuestro criterio de relevancia están supuestos todos nuestros valores**. La decisión de que una imagen del mundo es verdadera (o verdadera según nuestros conocimientos actuales o «tan verdadera como la que más») y las respuestas a las cuestiones relevantes (así como nuestra capacidad para resolverlas) revelan todo nuestro sistema de compromisos valorativos, sobre el cual descansan. Un ser sin valores tampoco tiene hechos.

El examen del más simple enunciado permite ver cómo nuestros criterios de relevancia implican valores, al menos indirectamente. Consideremos la oración «El gato está sobre la estera». Si alguien emitiera este juicio en un determinado contexto, emplearía recursos conceptuales — las nociones de «gato», «sobre» y «estera» — proporcionados por una determinada cultura, cuya presencia y ubicuidad revelan algo de los intereses y valores de esa cultura, y de casi todas las culturas. Tenemos la categoría «gato» porque consideramos significativa la división del mundo en animales y no animales, y además nos interesa a qué especie pertenece un animal dado. Es relevante que haya un gato sobre la estera, y no meramente una cosa. Tenemos la categoría «estera» porque consideramos significativa la división de las cosas inanimadas en *artificiales y no-artificiales*, y además **nos interesa la naturaleza y el propósito que tiene determinado artefacto**. Es relevante que el gato esté sobre la estera, y no únicamente sobre una cosa. Y tenemos la categoría «sobre» porque nos interesa las *relaciones espaciales*.

**PUTNAM, H.** (1988:165-166). *Razón, verdad e historia*. Madrid: Tecnos.

Demonos cuenta: hemos examinado el más trivial enunciado imaginable, «El gato está sobre la estera», y nos hemos encontrado con las presuposiciones que hacen que este enunciado sea relevante en ciertos contextos incluyen la significatividad de las categorías *animado/inanimado*, *propósito y espacio*. Para un espíritu que no esté dispuesto a considerar relevantes tales categorías «El gato está sobre la estera» sería una observación tan irracional como sería el enunciado «El número de objetos hexagonales de esta habitación es 76», proferido a mitad de un *tete-a-tete* entre jóvenes amantes.

No sólo se muestran en nuestras categorías (artefactos, nombres de especies, términos para las relaciones espaciales) muchos hechos generales con respecto a nuestro sistema de valores, sino que, como veíamos ... en nuestro uso las palabras clasificatorias específicas («amable», «egoísta») también se revelan nuestros valores más específicos (sensibilidad y compasión, por ejemplo). **Para repetirlo una vez más, nuestros criterios de relevancia revelan todo nuestro sistema de valores.**

**PUTNAM, H.** (1988:166). *Razón, verdad e historia*. Madrid: Tecnos.

La relevancia de esta discusión sobre la relevancia en relación a la pregunta que surgió en el capítulo precedente («¿Cuál es valor de la racionalidad?») es inmediata. Si la «racionalidad» es una capacidad (o mejor dicho, un sistema integrado de capacidades) que permite al que posee determinar qué preguntas son relevantes y qué respuestas justificadamente aceptables, entonces la racionalidad lleva el valor puesto. Pero no es necesario argumentar que esta concepción de la racionalidad está tan cargada de valor como la propia noción de relevancia.

... Una persona racional ... sería una persona que puede distinguir entre lo que puede y lo que no puede afirmarse *justificadamente*; lo que una persona elige considerar *interesante o importante o relevante* podría resultar pertinente a la hora de evaluar su carácter o su salud mental, pero no su racionalidad *cognitiva*.

Sin embargo, la aceptabilidad y la relevancia son interdependientes en cualquier contexto. El uso de cualquier palabra —sea ésta «bueno» o «consciente» o «rojo» o «magnético»— supone una historia, una tradición de observación, generalización, práctica y teoría.

**PUTNAM, H.** (1988:166-167). *Razón, verdad e historia*. Madrid: Tecnos.

También nos comprometemos con la actividad de *interpretar* esa tradición y de adaptarla a nuevos contextos, ampliándola y criticándola. Podemos interpretar la tradición de diversas formas, pero de ningún modo podemos aplicar una palabra si nos situamos totalmente fuera de la tradición a la que pertenecemos. Y el hecho de permanecer dentro de la tradición a la que pertenecemos afecta, con toda seguridad, a o que consideramos como «aceptabilidad racional». Si existe un método que pudiera utilizarse para verificar cualquier enunciado, sin importar qué conceptos contuviese, entonces sería factible mantener la separación propuesta entre la capacidad de verificar un enunciado y el dominio de un conjunto *relevante* de conceptos. Pero ya hemos visto que no hay razones para aceptar el mito del Único Método.

## **LA TEORÍA DE LOS COMPONENTES**

Nuestras intuiciones con respecto a la racionalidad parecen hallarse en conflicto, y ninguna filosofía parece reconciliarlas del todo. Por una parte, es sencillamente falso que *nunca* juzguemos los fines como racionales o irracionales; por otra, cuando nos enfrentamos con un caso como el del hipotético «nazi racional», no vemos cómo justificar el que desestimemos por irracional un sistema de fines tan inteligentemente elaborado y meditado, si bien lo encontramos moralmente repugnante.

**PUTNAM, H.** (1988:167). *Razón, verdad e historia*. Madrid: Tecnos.

Se ha sugerido que podemos resolver estos problemas del siguiente modo: si asumimos una nítida dicotomía hecho-valor, podemos justificar el que tachemos de irracional al hombre que se interesa sólo en saber el número de pelos que la gente tiene en la cabeza, sobre la base de que tiene una percepción inadecuada de los *hechos*... El nazi únicamente discrepa de nosotros respecto a los valores, y no por lo irracional. Quizá podamos tratar los casos intermedios siguiendo la línea sugerida por B. Williams. En particular, nuestro argumento contra el fetichista del método, esto es, que no podemos desechar la «evidencia observacional» de juicios descriptivos como «Juan es amable» sin caer en circularidad, puede contrastarse proponiendo la teoría que establece que el vocabulario descriptivo-moral del lenguaje ordinario tiene simultáneamente **dos «componentes de significado»**. Un componente es ficticio: hay ciertos estándares de amabilidad generalmente aceptados, y «Juan es amable» transmite la información de que Juan satisface esos estándares. Pero también hay un componente de significado *emotivo*: «Juan es amable» transmite una «actitud favorable» hacia cierto aspecto de la conducta de Juan. Se afirma que lo que puede ser racional es la aceptación del componente fáctico del enunciado «Juan es amable»; la aceptación del componente emotivo no es ni racional ni irracional.

**PUTNAM, H.** (1988:170). *Razón, verdad e historia*. Madrid: Tecnos.

**La idea de Saúl Kripke de verdades «metafísicamente necesarias» que han de ser aprendidas *empíricamente***, la idea de «verdades necesarias epistemológicamente contingentes» que penetró en la filosofía del lenguaje pocos años después de que yo introdujera la «identidad sintética de propiedades», amplía e ilumina el punto en el que estoy insistiendo. La observación de Kripke, aplicada al caso de la temperatura/energía cinética, es que si alguien describe un mundo lógicamente posible en el que la gente tiene sensaciones de calor y frío, en el que hay objetos que transmiten la sensación de calor y objetos que transmiten la de frío y en el que estas sensaciones se explican por un mecanismo distinto de la energía cinética molecular media, no es energía molecular media. Más bien diríamos que ha descrito un mundo en el que algún mecanismo diferente de la temperatura hace que ciertos objetos se perciban calientes y otros fríos. Una vez hemos aceptado el «enunciado de identidad sintética» según el cual la temperatura es energía cinético-molecular media (en el mundo actual) , no hay nada que cuente como un mundo posible en el que la temperatura no es energía cinética molecular media.



**PUTNAM, H.** (1988:170). *Razón, verdad e historia*. Madrid: Tecnos.

**Tradicionalmente se ha llamado «necesario» a un enunciado que es verdadero en cualquier mundo posible. Y «esencial» a una propiedad que algo tiene en todo el mundo posible.** Saúl Kripke está afirmando, sirviéndose de esta terminología tradicional, que el enunciado «La temperatura es energía cinético molecular media» **es una *verdad necesaria*, aun cuando no podemos saberlo a priori.** El enunciado es empírico, pero necesario. O, por decirlo con otras palabras, ser energía cinético molecular media es una propiedad esencial de la temperatura: hemos descubierto la esencia de la temperatura mediante investigación empírica. Estas ideas de Kripke han tenido gran impacto en la filosofía del lenguaje, la metafísica y la filosofía de las matemáticas; aplicadas al argumento de Moore son devastadoras.

**PUTNAM, H.** (1988:170-171). *Razón, verdad e historia*. Madrid: Tecnos.

Moore argumentaba que a partir del hecho de que (1) *sólo puede ser contingentemente falso*, que ser P (para alguna adecuada propiedad natural P) no puede ser una propiedad esencial de la *bondad*; y esto es lo que la nueva teoría de la necesidad impide. Todo lo que se puede inferir válidamente a partir del hecho de que (1) no es autocontradictorio es que «bueno» no es sinónimo de «conducente a la máxima utilidad» (no es sinónimo de P, para cualquier término P pertenece a la versión fisicalista del mundo). De esta no-sinonimia de palabras no se sigue nada acerca de la no-identidad de propiedades. No se sigue nada con respecto a la esencia de la bondad.

Ruth Anna Putnam ha señalado que existe un argumento común en favor de que la bondad no puede ser una propiedad natural que tampoco resuelve el problema: el argumento de que «X es bueno» tiene fuerza emotiva, expresa una «actitud favorable», etc.

Ruth Anna Putnam señala que este argumento no tiene en cuenta que muchos predicados adquieren con naturalidad cierta fuerza emotiva. «Babeó la comida sobre su camisa» es un enunciado que tiene literalmente en nuestra cultura una fuerza marcadamente negativa, a pesar de ser literalmente una descripción. **Cualquier palabra que represente algo que la gente de una cultura valore (o devalúe) tenderá a adquirir una fuerza emotiva.**

**PUTNAM, H.** (1988:171). *Razón, verdad e historia*. Madrid: Tecnos.

La palabra «bueno», en su sentido moral, se aplica a muchas cosas ... algunas pueden ser valoradas por su naturaleza: puede formar parte de ese mismo estado de ánimo el que uno valore encontrarse en él... Supongamos que «bueno» se definiese de modo que las cosas que se valoran por su naturaleza, y no existe ninguna buena razón para *devaluarlas* (como la hay para devaluar ciertos estados de ánimo inducidos por las drogas, los cuales a veces se valoran por su naturaleza), cuenten como «buenas».

Entonces, es de esperar que el enunciado que afirma que algo es bueno tenga una «fuerza emocional» positiva debido a la naturaleza de esa propiedad. Aun cuando uno no sea cuensecuencialista (es decir, alguien que piense que cualquier cosa que produzca consecuencias suficientemente buenas, es buena), no hay duda de que a razón más común para llamar buena a una acción es que tenga buenas consecuencias, entre las cuales podría estar el que promoviese estados o situaciones que se valoran por su naturaleza; una vez más, la misma naturaleza de la propiedad explica por qué la descripción llega a tener fuerza emotiva «favorable».

**PUTNAM, H.** (1988:170). *Razón, verdad e historia*. Madrid: Tecnos.

Makie defiende que la bondad es una propiedad ontológicamente «extraña» introduciendo como premisa la suposición de que no es posible *saber* que algo es bueno sin tener una actitud «favorable» hacia ese algo. Y ello equivale a suponer en *emotivismo* para probar el emotivismo... ¿Acaso nunca hemos oído decir: «Sé que está mal hacerlo, pero no me importa?». Como ha señalado Philippa Foot, uno puede rechazar los llamamiento a la moralidad diciendo: «Yo no aspiro a ser un hombre bueno». **La mera posibilidad de estas preferencias muestra que aunque exista en realidad una diferencia en el uso descriptivo del lenguaje y el uso prescriptivo o aprobatorio, tal diferencia no está en función del vocabulario.** Las palabras «descriptivas» pueden usarse para alabar o denostar... y las palabras «evaluativas» pueden usarse para describir y explicar.

**PUTNAM, H.** (1988:172-173). *Razón, verdad e historia*. Madrid: Tecnos.

La profesora Putnam señala que el argumento de Makie acierta en una cosa, sin embargo. **Es innegable que algunas expresiones morales llevan incorporada cierta orientación hacia la acción.** «Yo debería», «correcto» y «deber» son los principales ejemplos de expresiones guía-para-la-acción. Como ella señala, el problema del «es debe» no coincide con el problema del «hecho/valor». «No tengo intención de hacer lo que debería hacer» nos suena más extraño que «No tengo la intención de ser un hombre bueno» (y «No tengo la intención de hacer lo que debo» parece un disparate).

Makie señala que ninguna propiedad física tiene incorporada una conexión con la acción (o con la aprobación de la acción) y concluye que «ser lo correcto a realizar», etc., son «ontológicamente extrañas». Pero su argumento, además de depender de la suposición de que la versión fisicalista del mundo es la teoría verdadera, prueba demasiado. Pues algunos predicados *epistémicos* (por ejemplo «racionalmente aceptable», «creencia justificada») son también **guía-para-la-acción**... Alguien puede decir «Está bien hacer X» y «Hay bastante evidencia de que y» y no comprometerse a hacer o prescribir X o a aceptar Y; pero si alguien dice «En esta situación es correcto realizar la acción X» o «Creo que X está completamente justificado» (en el caso de las creencias) tiene a característica de ser guía-para-la-acción, como la tiene «correcto», en sentido moral. Si admitimos la teoría de conjuntos como parte de una versión naturalista del mundo). Si la relación de referencia fuera una de éstas, entonces ese hecho sería en sí mismo un hecho metafísico último de muy extraña índole,

**PUTNAM, H.** (1988:173). *Razón, verdad e historia*. Madrid: Tecnos.

Lo que convierte a este hecho en algo tan extraño es que hemos incorporado cierta *naturalidad*, cierta inconsistencia o carencia de deliberación (*mindlessness*) en nuestra misma concepción de la naturaleza. **Se supone que la naturaleza no tiene intenciones, intereses o puntos de vista.** Dado esto por cierto, **¿cómo podría estar seleccionada metafísicamente una relación de referencia admisible?**

Esa misma inconsistencia de la naturaleza hace parecer «extraños» los predicados guía-para-la-acción «Es correcto» y «Es una creencia justificada». **Sería extraño que una propiedad fisicalista P fuera idéntica a la corrección moral o a la justificación epistémológica** —precisamente por la misma razón por la que sería «extraño» que la referencia fuera una relación fisicalista.—. **Sería como si la naturaleza tuviese valores, en el caso moral, o intenciones referenciales, en el caso semántico.**

Por esta razón, pienso que Moore estaba en lo cierto (aun cuando sus argumentos no son aceptables) al sostener que «bueno», «correcto» (y también »creencia justificada«, »refiere« y »verdadero«) no son idénticas propiedades y relaciones fisicalistas. Y lo que esto muestra no es que no existan la bondad, la corrección, la justificación epistémológica y la verdad, sino que el naturalismo monista (o «fisicalismo») es una filosofía inadecuada.

**PUTNAM, H.** (1988:176). *Razón, verdad e historia*. Madrid: Tecnos.

Hemos investigado la cuestión de si los «juicios de valor» pueden apoyarse racionalmente. Hemos visto que las diversas respuestas negativas descansan en suposiciones filosóficas dudosas: que la misma racionalidad sólo es buena para la «predicción» o para obtener «consenso», o que sólo hay un método para llegar a la verdad (a veces se afirma que el único criterio de verdad lo constituyen la predicción y el consenso), o que las propiedades de valor son «ontológicamente extrañas». **La posición que he defendido es que cualquier elección de esquema conceptual presupone valores**, y la elección de un esquema conceptual para describir las relaciones interpersonales y los hechos sociales, por no mencionar la reflexión sobre el plan de vida mismo, implica, entre otras cosas, los valores morales que uno mantiene. No puede elegirse un esquema que simplemente «copie» los hechos, ya que *ningún esquema conceptual es una mera «copia del mundo»*. **El contenido de la misma noción de verdad depende de los criterios de aceptabilidad, y éstos, a su vez, presuponen nuestros valores, sobre los que descansan. Expresándolo esquemática y brevemente: la teoría de la verdad presupone la teoría de la racionalidad, que a su vez presupone nuestra teoría de lo bueno.**

La «teoría de lo bueno», sin embargo, no es sólo programática, sino que depende de suposiciones acerca de la naturaleza humana, de la sociedad y del universo (incluyendo suposiciones teológicas y metafísicas). Hemos tenido que revisar nuestra teoría de lo bueno (tal como está) una y otra vez, según va aumentando nuestro conocimiento y va cambiando nuestra cosmovisión.

**PUTNAM, H.** (1988:176-177). *Razón, verdad e historia*. Madrid: Tecnos.

Creo que es evidente que no existe algo como un «fundamento» para la concepción que defiende. Y en este punto la gente se inquieta: ¿no nos hallamos muy próximos al punto de vista de que no hay diferencia entre «justificado» y «justificado para mis conocimientos» (una especie de solipsismo?)...

No estamos atrapados en los infiernos solipsistas individuales, sino invitados a tomar parte en un diálogo genuinamente humano, un diálogo que combine la colectividad con la responsabilidad individual.

¿Tiene este diálogo un término ideal? ¿Hay una concepción verdadera de la racionalidad, una moralidad verdadera, aun cuando todo lo que tenemos son nuestras *concepciones de éstas*?

En este punto, las opiniones de filósofos, como las de los demás, se dividen. Richard Rorty, en su discurso presidencial a la *American Philosophical Association*, optó con firmeza por un punto de vista de que sólo existe el diálogo: no puede postularse ningún fin ideal, ni tampoco sería necesario. Pero la afirmación de que «sólo existe el diálogo» ... ¿difiere en algo del relativismo...? El mismo hecho de que hablemos de nuestras diferentes concepciones como diferentes concepciones de la *racionalidad* postula un *Genz-begriff*, un *concepto límite de verdad ideal*.