

PRIMERA PARTE
TEORÍA DEL CONOCIMIENTO

En Miller, D. (1995). *Popper, Escritos Selectos*. México: Fondo de Cultura Económica. Pp 10-70

1. LOS COMIENZOS DEL RACIONALISMO (1958)

I. REMONTÉMONOS A LOS PRESOCRÁTICOS

"REMONTARSE hasta Matusalén" era un programa progresivo, comparado con "Remontarse hasta Tales", o "Remontarse hasta Anaximandro": lo que Shaw nos ofreció era una expectativa de vida mejorada —algo que estaba en el aire, por lo menos cuando él lo escribió—. Me temo que no tengo que ofrecer nada que "esté en el aire" ahora; pues a lo que quiero remontarme es a la sencilla y llana *racionalidad* de los presocráticos. ¿En qué estriba esta tan llevada y traída "racionalidad" de los presocráticos? La sencillez y la osadía de sus preguntas es parte de ella, pero mi tesis es que el punto decisivo lo constituye la actitud crítica que, como intentaré demostrar, se desarrolló por vez primera en la escuela jónica.

Las preguntas que los presocráticos trataban de contestar eran primordialmente de índole cosmológica, pero también había preguntas referentes a la teoría del conocimiento. Estoy convencido de que la filosofía debe volver a la cosmología y a formular una sencilla teoría del conocimiento. Existe por lo menos un problema filosófico en que todos los pensadores están interesados: el de entender el mundo en que vivimos; y por ende, en entendernos a nosotros mismos (que somos parte de ese mundo) y nuestro conocimiento de él. Yo creo que toda ciencia es cosmología, y para mí el interés de la filosofía, no menos que el de la ciencia, radica únicamente en este osado intento de ampliar nuestro conocimiento del mundo y la teoría de nuestro conocimiento del mundo. Me interesa Wittgenstein, por ejemplo, no por su filosofía lingüística, sino porque su *Tractatus* era un tratado de cosmología (aunque burdo) y porque su teoría del conocimiento estaba estrechamente vinculada a su cosmología.

Para mí, tanto la ciencia como la filosofía pierden todo su atractivo cuando se apartan de esa búsqueda —cuando se convierten en especialismos y dejan de ver, y de maravillarse por, los acertijos de nuestro mundo—. La especialización puede ser una gran tentación para el científico. Para el filósofo constituye un pecado mortal.

II. LA TRADICIÓN DE LA DISCUSIÓN CRÍTICA

La más antigua historia de la filosofía griega, especialmente la historia que va de Tales a Platón, es un relato espléndido. Es casi demasiado bueno para ser verdad. En cada generación encontramos por lo menos una nueva filosofía, una nueva cosmología de pasmosa originalidad y profundidad. ¿Cómo fue esto posible? Por supuesto, no podemos explicar ni la originalidad ni el genio. Pero podemos intentar arrojar alguna luz al respecto. ¿Cuál era el secreto de los antiguos filósofos? Sugiero que fue una *tradición: la tradición de la discusión crítica*.

Trataré de plantear el problema de modo más claro. En todas o en casi todas las civilizaciones encontramos algo que podemos llamar enseñanza religiosa y cosmológica, y en muchas sociedades encontramos escuelas. Ahora las escuelas, especialmente las primitivas, tienen, al parecer, una estructura y funciones características. Lejos de ser lugares dedicados a la discusión crítica, se consagran a la tarea de enseñar determinada doctrina y a conservarla, pura y sin cambios. Es tarea de toda escuela transmitir la tradición, la doctrina de su fundador, de su primer maestro, a la siguiente generación, y para este fin lo más importante radica en conservar inviolada esa doctrina. Una escuela de esta clase jamás admite una nueva idea. Las ideas nuevas son herejías, y redundan en cismas: si un integrante de la escuela intenta cambiar la doctrina, es expulsado como un hereje. Pero el hereje arguye, generalmente, que la suya es la verdadera doctrina del fundador. Así, ni siquiera el inventor reconoce que ha introducido una invención; más bien está convencido de que está volviendo a la verdadera ortodoxia, la cual, de alguna forma, se ha pervertido.

De esta manera, todo cambio a la doctrina —si acaso hay alguno— es un cambio subrepticio. Todos los cambios se presentan como reformulaciones de los verdaderos conceptos que ha emitido el maestro, de sus propias palabras, del significado que él les dio, de sus verdaderas intenciones.

Está claro que en una escuela de esta clase no debemos esperar encontrar una historia de las ideas, ni siquiera material para tal historia. Porque las nuevas ideas no se reconocen como nuevas. Todo se adscribe al maestro. Lo único que podríamos reconstruir es una historia de los cismas, y acaso la historia de la defensa de ciertas doctrinas contra los herejes.

Por supuesto, no puede haber ninguna discusión racional en una escuela de esta clase. Puede haber argumentos en contra de los disidentes y herejes, o en contra de algunas escuelas opuestas. Pero, principalmente, la doctrina de esa escuela se defiende con aseveraciones, dogmas y condenaciones, más que con argumentos.

El gran ejemplo de una escuela de esta índole entre las escuelas filosóficas de Grecia es la escuela italiana que fundó Pitágoras. Comparada con la jónica o con la eleática, tenía el carácter de una orden religiosa, con un modo de vida característico y una doctrina secreta. La historia de que un integrante de esta escuela, Hipaso de Metaponto, fue ahogado en el mar por haber revelado el secreto de la irracionalidad de ciertas raíces cuadradas, es característico de la atmósfera que rodeaba a la escuela pitagórica, sea o no verídico este relato.

Pero entre las escuelas filosóficas griegas los primeros pitagóricos fueron una excepción. Dejándolos aparte, podríamos afirmar que el carácter de la filosofía griega, y de sus escuelas filosóficas, es notablemente diferente del tipo dogmático de escuela que aquí se describe. [En el trabajo número 18] Ilustro esto con un ejemplo: *la historia del problema del cambio es la historia de un debate crítico, de una discusión racional*. Las nuevas ideas se proponen como tales, y surgen como resultado de la crítica abierta. Hay pocos cambios subrepticios, si acaso los hay. Y en vez del anonimato, encontramos una historia de las ideas y de sus creadores.

Estamos aquí ante un fenómeno único en su género, estrechamente vinculado con las asombrosas libertad y creatividad de la filosofía griega. ¿Cómo explicar este fenómeno? *Lo que tenemos que explicar es el surgimiento de una tradición*. Es una tradición que permite o anima las discusiones críticas entre varias escuelas y, por más asombroso que parezca, en el seno de una misma escuela. Pues en ninguna parte, fuera de la escuela pitagórica, encontramos una escuela consagrada a conservar una doctrina. En vez de ello, encontramos cambios, nuevas ideas, modificaciones y rotundas críticas al maestro.

(En Parménides incluso encontramos, en fecha temprana, el más notable fenómeno: el de un filósofo que preconiza *dos* doctrinas, una que afirma ser verdadera, y otra que él mismo describe como falsa. Sin embargo, no convierte a la doctrina falsa simplemente en un objeto de condenación o de crítica; más bien la presenta como la mejor explicación posible de la engañosa opinión de los mortales y del mundo de la mera apariencia, la mejor explicación que un mortal puede dar.)

¿Cuándo y en dónde se fundó esta tradición crítica? He aquí un problema que merece meditarse seriamente. Una cosa es cierta: Jenófanes, que llevó la tradición jónica a Elea, estaba del todo consciente de que su propia enseñanza era puramente conjetural, y de que vendrían otros que sabrían más que él. Volveré a ocuparme de este punto más adelante, en la sección m.

Si buscamos los primeros indicios de esta nueva actitud crítica, de esta nueva libertad de pensamiento, llegamos hasta la crítica que hace Anaxi-

mandro al pensamiento de Tales. [Véase, más adelante, el escrito 18.] He aquí un hecho extraordinario: Anaximandro critica a su maestro y pariente, uno de los Siete Sabios, el fundador de la escuela jónica. Anaximandro era entonces, según la tradición, sólo unos 14 años más joven que Tales, y debió de desarrollar su crítica y sus nuevas ideas en vida de su maestro. (Parece que ambos murieron con pocos años de diferencia uno del otro.) Pero no encontramos en las fuentes de información ni una huella de disensión, de alguna querella o de algún cisma entre estos dos filósofos.

Esto sugiere, en mi opinión, que fue Tales quien fundó la nueva tradición de libertad de pensamiento —basada en una nueva relación entre maestro y discípulo— y que así creó un nuevo tipo de escuela, del todo diferente de la pitagórica. Parece haber sido capaz de tolerar la crítica. Y lo que es más, parece haber creado la tradición de que uno debería tolerar la crítica.

Pero me gusta pensar que incluso creó algo más: apenas puedo imaginar esta relación entre maestro y discípulo en la que el maestro sólo tolera la crítica, sin animarla activamente. No considero posible que un discípulo a quien se preparaba en la actitud dogmática se atreviera a criticar el dogma (y mucho menos el de un famoso sabio) y a expresar abiertamente su crítica. Y me parece una explicación más fácil y sencilla de suponer que el maestro alentaba la actitud crítica —posiblemente no desde el principio de la enseñanza, sino sólo después de haberle impresionado la pertinencia de algunas preguntas que le planteaba el discípulo, aun sin intención crítica—.

Como haya sido en realidad, la conjetura de que Tales alentaba activamente la crítica en sus discípulos explicaría el hecho de que la actitud crítica hacia la doctrina del maestro se volvió parte de la tradición de la escuela jónica. Me gusta pensar que Tales fue el primer maestro que dijo a sus discípulos: "Así es como veo yo las cosas; como creo que son las cosas. Traten de mejorar lo que les enseño". (A quienes crean que es "no histórico" atribuir esta actitud adogmática a Tales, les vuelvo a recordar el hecho de que sólo dos generaciones después encontramos una actitud similar consciente y claramente formulada en los fragmentos de Jenófanes.) De cualquier manera, el hecho histórico es que la escuela jónica fue la primera en que los discípulos criticaban a sus maestros, generación tras generación. No puede haber muchas dudas en el sentido de que la tradición griega de la crítica filosófica tuvo su fuente principal en Jonia.

Fue aquélla una importantísima innovación. Significó el rompimiento con la tradición dogmática que permitía sólo *una* doctrina, y la introducción, en vez de ella, de una tradición que admite una *pluralidad* de

doctrinas, las cuales tratan de aproximarse a la verdad por medio de la discusión crítica.

Así, esta actitud tiene como corolario, casi por necesidad, el darnos cuenta de que nuestros intentos de ver y encontrar la verdad no son definitivos, sino que siempre están abiertos a mejorarse; de que nuestro conocimiento, nuestra doctrina, es siempre conjetural; de que consiste en supuestos, o hipótesis, más que en verdades definitivas o certeras; y de que la crítica y la discusión crítica son los únicos medios de que disponemos para acercarnos más a la verdad. De esta manera, nos lleva a abrazar la tradición de las conjeturas atrevidas y de la crítica libre, la tradición que creó la actitud racional o científica y, con ella, nuestra civilización occidental, la única civilización basada en la ciencia (aunque, por supuesto, no únicamente en la ciencia).

En esta tradición racionalista, los cambios radicales de doctrina no están prohibidos. Al contrario, se alienta la innovación, y esta innovación se considera un éxito, una mejora, si está basada en los resultados de una discusión crítica de las doctrinas predecesoras. La osadía misma de una innovación es objeto de admiración; porque es posible controlarla con la severidad de su examen crítico. Por esta razón los cambios de doctrina, lejos de hacerse de manera subrepticia, se transmiten tradicionalmente junto con las viejas doctrinas y los nombres de los innovadores. Y el material para la historia de las ideas se convierte en parte de la tradición de la escuela.

Que yo sepa, la tradición crítica o racionalista se inventó una sola vez. Se perdió al cabo de dos o tres siglos, quizá debido al surgimiento de la doctrina aristotélica del *epistémé*, del conocimiento cierto y demostrable (desarrollo de la distinción eleática y heraclitana entre la verdad certera y la simple conjetura). Esta tradición se redescubrió y revivió conscientemente en el Renacimiento, especialmente con la obra de Galileo Galilei.

III. EL RACIONALISMO CRÍTICO

Ahora me ocuparé de mi aseveración central. Consiste en esto: la tradición racionalista, la tradición de la discusión crítica, representa el único medio practicable para ampliar nuestro conocimiento —el conocimiento conjetural o hipotético, por supuesto—. No existe otra manera de hacerlo. Más especialmente, no hay camino que parta de la observación o de la experimentación. En el desarrollo de la ciencia, las observaciones y los experimentos sólo desempeñan el papel de argumentos de crítica. Y juegan este papel junto con otro, el de argumentos no surgidos de la observación. Es un papel importante; pero el significado de las observa-

ciones y de los experimentos depende *enteramente* de si pueden utilizarse o no para *criticar teorías*.

Según la teoría del conocimiento aquí esbozada, existen, en última instancia, sólo dos maneras en que las teorías pueden ser superiores a otras: pueden explicar más, y pueden probarse mejor; esto es, pueden ser discutidas más cabal y críticamente, a la luz de todo cuanto sabemos, de todas las objeciones en que podamos pensar, y especialmente también a la luz de las pruebas de observación y de experimentación que se hayan diseñado con el fin de criticar la teoría.

Existe sólo un elemento de racionalidad en nuestros intentos por conocer el mundo: el examen crítico de nuestras teorías. Estas teorías en sí mismas son conjetura. No sabemos: sólo conjeturamos. Si me preguntara usted: "¿Cómo sabe...?", mi respuesta sería: "No sé; sólo propongo una suposición. Si está usted interesado en mi problema, me complacerá mucho que critique mi suposición, y si expresa usted contrapropuestas, yo, a mi vez, trataré de criticarlas".

Estoy convencido de que esto constituye la verdadera teoría del conocimiento (que deseo someter a la crítica del lector): la verdadera descripción de una práctica que surgió en Jonia y que se ha incorporado a la ciencia moderna (aunque hay muchos científicos que aún creen en el mito baconiano de la inducción): la teoría de que el conocimiento se desarrolla por medio de *conjeturas y refutaciones*.

Dos de los más grandes hombres que vieron con toda claridad que no existía tal procedimiento inductivo, y que entendieron bien lo que yo considero la verdadera teoría del conocimiento, fueron Galileo y Einstein. Sin embargo, los antiguos también lo supieron. Por increíble que parezca, descubrimos un claro reconocimiento y una clara formulación de esta teoría del conocimiento racional casi inmediatamente después de empezar a practicarse la discusión crítica. Los más antiguos fragmentos que han llegado hasta nosotros en este campo son los de Jenófanes. Presentaré aquí cinco de ellos, en un orden que sugiere que fueron la fuerza de su ataque y la gravedad de sus problemas los que le hicieron tomar conciencia de que todo nuestro conocimiento era conjetural, pero que, sin embargo, mediante la búsqueda del conocimiento "que sea el mejor", lo encontraremos con el tiempo. He aquí los cinco fragmentos de los escritos de Jenófanes:

Los etíopes dicen que sus dioses son de nariz achatada y negros,
mientras que los tracios dicen que los suyos tienen ojos azules y pelo rojo.

Pero si el ganado, o los caballos, o los leones, tuvieran manos y supieran dibujar

y pudieran esculpir, como los hombres, los caballos dibujarían a sus dioses con forma de caballos, y el ganado, con forma de ganado, y luego cada cual daría forma a los cuerpos de los dioses a semejanza, cada clase de ser, de su propia forma.

Los dioses no nos revelaron, desde el principio,
todas las cosas; pero con el paso del tiempo,
mediante la búsqueda, los hombres descubren lo que es mejor...

Estas cosas son, conjeturamos, como la verdad.

Pero en cuanto a la verdad cierta, ningún hombre la ha conocido,
y no la conoceremos; ni la verdad acerca de los dioses,
ni tampoco acerca de todas las cosas de las que hablo.
Y hasta si por casualidad le fuera dado al hombre pronunciar
la verdad definitiva, él mismo no lo sabría:
porque todo no es sino una maraña de suposiciones.

Para demostrar que Jenófanes no estaba solo en esto, puedo reproducir aquí dos de las enseñanzas de Heráclito que ya he citado en un contexto diferente. Ambas expresan el carácter conjetural del conocimiento humano, y la segunda se refiere a su osadía, a la necesidad de anticipar atrevidamente lo que no sabemos:

No está en la naturaleza ni en el carácter del hombre poseer el verdadero conocimiento, aunque éste pertenece a la naturaleza divina. [...] Quien no espere lo inesperado, no lo detectará: para éste, seguirá siendo indetectable e inabordable.

Mi última cita es **una**, muy famosa, de Demócrito:

Pero, de hecho, no sabemos nada por haberlo visto; porque la verdad está oculta en las profundidades.

Así es como la actitud crítica de los presocráticos anunció y preparó el racionalismo ético de Sócrates: su convicción de que la búsqueda de la verdad mediante la discusión crítica era un modo de vida... el mejor que él conoció.¹

¹ Los fragmentos que se citan aquí son, respectivamente, de Jenófanes, B 16, 15, 18, 35 y 34; de Heráclito, B 78 y 18, y de Demócrito, B 117, en H. Diels y W. Krantz, *Die Fragmente der Vorsokratiker*, 5ª edición, 1964. [Las traducciones al inglés son de Popper mismo. Otra traducción de todos los fragmentos presocráticos, en el mismo orden de la obra de Diels y Krantz, puede verse en Kathleen Freeman, *Ancilla to The Pre-Socratic Philosophers*, 1948.]

2. LA DEFENSA DEL RACIONALISMO (1945)

LA QUERELLA entre el racionalismo y el irracionalismo data de muy antiguo. Aunque la filosofía griega indudablemente comenzó como una empresa racionalista, hubo trazas de misticismo aun en sus inicios. Es la nostalgia de la perdida unidad y del resguardo seguro del tribalismo la que se expresa en estos elementos místicos con un enfoque fundamentalmente racional.¹ Un abierto conflicto entre el racionalismo y el irracionalismo estalló por vez primera en la Edad Media, como la oposición entre el escolasticismo y el misticismo. En los siglos xvn, xvm y xix, cuando la marea del racionalismo, del intelectualismo y del "materialismo" iba en aumento, los irracionales tuvieron que darle cierta atención, argumentando en su contra; y al exponer sus limitaciones, y las inmoderadas pretensiones y los peligros del seudorracionalismo (que ellos no distinguían del racionalismo en el sentido actual), algunos de estos críticos, sobre todo Burke, se ganaron la gratitud de los verdaderos racionalistas. Pero luego se revirtió la marea, y "alusiones profundamente significantes [...] y alegorías" (como lo expresó Kant) llegaron a estar de moda entre los pensadores de la época. Un irracionalismo oracular ha establecido (especialmente con Bergson y la mayoría de los

¹ Véase *La sociedad abierta y sus enemigos*, capítulo x, especialmente las notas 38-41 y el texto mismo.

En Pitágoras, Heráclito, Parménides y Platón, los elementos místicos y racionalistas están mezclados. Platón especialmente, pese a todo el énfasis que pone en "la razón", incorporó a su filosofía tan considerable mezcla de irracionalismo, que casi desechó el racionalismo que heredó de Sócrates. Esto permitió a los neoplatónicos basar su misticismo en Platón; y la mayor parte del misticismo subsecuente se remonta hasta estas fuentes.

Quizá sea accidental, pero en todo caso es notable comprobar que todavía existe una frontera cultural entre la Europa occidental y las regiones de la Europa central que no estuvieron sujetas a la administración del Imperio Romano de Augusto, y que no gozaron de los beneficios de la *Pax Romana*, es decir, de la civilización romana. Las mismas regiones "bárbaras" se muestran particularmente inclinadas a la influencia del misticismo, aunque ellas no lo inventaron. Bernardo de Clairvaux obtuvo sus mayores éxitos en Alemania, donde florecieron posteriormente Eckhart y su escuela, y también Boehme.

Mucho tiempo después, Spinoza, que intentó combinar el intelectualismo cartesiano con las tendencias místicas, redescubrió la teoría de la intuición mística intelectual, la cual, pese a la fuerte oposición de Kant, redundó en el surgimiento poskantiano del "idealismo", hasta Fichte, Schelling y Hegel. Casi todo el irracionalismo moderno se remonta hasta este último, como se ha indicado brevemente en el capítulo 12 de *La sociedad abierta y sus enemigos*.

filósofos e intelectuales alemanes) el hábito de pasar por alto o al menos deplorar la existencia de seres tan "inferiores" como los racionalistas. Para tales pensadores, los racionalistas —o los "materialistas", como a menudo los catalogan—, y especialmente el científico racionalista, son los pobres de espíritu, que se dedican a actividades sin alma y en gran medida mecánicas, y que no captan en absoluto los más profundos problemas del destino humano y de su filosofía. Y los racionalistas, por su parte, reaccionan a la recíproca, desautorizando al irracionalismo como una rotunda tontería. Nunca como ahora ha sido esta ruptura tan radical. Y el rompimiento de relaciones diplomáticas entre los filósofos probó su importancia cuando le siguió la ruptura de relaciones diplomáticas entre los Estados.

En esta querella, estoy enteramente de parte del racionalismo. Esto es tan patente que, incluso cuando siento que el racionalismo ha ido demasiado lejos, sigo simpatizando con él, pues sostengo que un exceso en esta dirección (en la medida en que excluyamos la falta de humildad intelectual del seudorracionalismo de Platón) es en realidad inocuo, si lo comparamos con un exceso en la otra dirección. Opino que la única manera en la que el racionalismo excesivo puede ser dañino es que tienda a minar su propia posición y, con ello, acarrear una reacción irracionalista. Es sólo este peligro el que me induce a examinar más detenidamente las pretensiones de un racionalismo excesivo y abogar por un racionalismo autocrítico y modesto que reconozca tener ciertas limitaciones. En consecuencia, en lo que sigue distinguiré entre dos posiciones racionalistas, a las que llamaré "racionalismo crítico" y "racionalismo acrítico" o "racionalismo total".

El racionalismo acrítico o total puede describirse como la actitud de la persona que dice: "No estoy dispuesto a aceptar nada que no pueda defenderse por medio de un argumento o de la experiencia". También podemos expresarlo en la forma del principio según el cual toda suposición que no pueda apoyarse, ya sea en la argumentación o en la experiencia, tiene que descartarse.² En este punto es fácil darse cuenta de que este principio de un racionalismo acrítico es incongruente, dado que, como a su vez no puede apoyarse en argumentos válidos o en la experiencia, implica que él mismo tendría que ser descartado. (Esto es análogo a la paradoja del mentiroso,³ es decir, a una frase que afirma su propia

² Digo "descartarse" para abarcar los puntos de vista siguientes: 1) que tal suposición sería falsa; 2) que sería acientífica (o impermisible), aunque quizá resultara accidentalmente verdadera; 3) que sería "carente de sentido o sin significado", por ejemplo, en el sentido del *Tractatus Logico-Philosophicus* de Wittgenstein [véase la nota 17 al escrito 6, y la nota 4(2), más adelante].

³ En ésta, y en la nota siguiente, se harán algunas observaciones sobre las paradojas, especialmente sobre la *paradoja del mentiroso*. Al presentar estas observaciones, conviene

falsedad.) El racionalismo acrítico es, por tanto, lógicamente insostenible; y puesto que un argumento puramente lógico puede demostrarlo, es posible derrotarlo con la misma arma que ha elegido: el argumento.

Esta crítica puede ser generalizada. Como todo argumento debe proceder a partir de suposiciones, resulta del todo imposible exigir que todo

decir que las llamadas paradojas "lógicas" o "semánticas" ya no son sólo juegos de palabras para los lógicos. No sólo han probado que son importantes para el desarrollo de las matemáticas, sino que también tienen importancia en otros campos del pensamiento. Hay una conexión definida entre estas paradojas y problemas tales como la *paradoja de la libertad* [véanse las notas 4 y 6 al escrito 25, y la sección III del escrito 26], que es de considerable importancia en filosofía política. En el punto 4) de esta nota se demostrará brevemente que las diversas *paradojas de la soberanía* son muy semejantes a la paradoja del mentiroso. Sobre los métodos modernos para resolver estas paradojas (o quizá, mejor dicho, para construir lenguajes en que no se presenten las paradojas) no haré aquí ningún comentario, pues nos llevarían muy lejos de nuestro tema.

1) La *paradoja del mentiroso* puede formularse de muchas maneras. Una de ellas es ésta: Supongamos que alguien dice un día: "Todo lo que digo hoy es mentira", o, más precisamente: "Todas las afirmaciones que hago hoy son falsas", y que no dice nada más todo ese día. Ahora bien, si nos preguntamos si este hombre dijo la verdad, he aquí lo que descubrimos. Si empezamos con la suposición de que lo que dijo era verdad, entonces llegamos a la conclusión, considerando *lo que* dijo, de que debió ser falso. Y si partimos de la suposición de que lo que dijo era falso, entonces debemos concluir, considerando *lo que* dijo, que fue verdad.

2) A las paradojas se les llama a veces "contradicciones". Pero esto quizá sea un tanto desorientador. Una contradicción ordinaria (o algo que se contradice a sí mismo) es simplemente una afirmación falsa desde el punto de vista lógico, como: "Platón era feliz ayer y no era feliz ayer". Si suponemos que tal enunciado es falso, no se suscita ninguna dificultad. Pero de una paradoja no podemos suponer ni que es verdadera *ni que es falsa*, sin meternos en dificultades.

3) Hay, por supuesto, afirmaciones estrechamente relacionadas con las paradojas, pero que sólo son, estrictamente, contradicciones de sí mismas. Veamos, por ejemplo, la aseveración "Todas las afirmaciones son falsas". Si suponemos que esta aseveración es verdadera, llegaremos a la conclusión, considerando *lo que* dice, de que es falsa. Pero si suponemos que es falsa, entonces salimos de la dificultad; porque entonces esta suposición nos hará concluir que no todas las afirmaciones son falsas. O, en otras palabras, que existen algunas afirmaciones —por lo menos una— que son verdaderas. Y este resultado es inocuo; porque no implica que nuestra afirmación originaria sea una de las verdaderas. (Esto no implica que podamos, de hecho, construir un lenguaje *libre de paradojas* en el que pueda formularse: "Todas las afirmaciones son falsas" o "Todas las afirmaciones son verdaderas".)

A pesar de que esta afirmación, "Todas las proposiciones son falsas", no es en realidad una paradoja, puede llamarse, por cortesía, "una forma de la paradoja del mentiroso", por su obvia semejanza con ella; y ciertamente, la antigua formulación griega de esta paradoja (Epiménides el Cretense dice "Todos los cretenses mienten siempre") es, en esta terminología, más bien "una forma de la paradoja del mentiroso", es decir, una contradicción, más que una paradoja. (Véase también la siguiente nota.)

4) Ahora demostraré brevemente la semejanza entre la paradoja del mentiroso y las diversas *paradojas de la soberanía*; por ejemplo, el principio de que los mejores o los más sabios o la mayoría deberían gobernar.

C. H. Langford ha descrito varias maneras de expresar la paradoja del mentiroso; entre ellas, la siguiente. Consideremos dos afirmaciones, que hacen respectivamente dos personas, A y B.

A dice: "Lo que dice B es verdad".

presupuesto o suposición deba basarse en argumentos válidos. La exigencia de muchos filósofos de que deberíamos empezar sin hacer suposiciones de ninguna especie y de que nunca deberíamos suponer nada sin "razón suficiente", e incluso la más débil exigencia de que deberíamos empezar con un muy pequeño conjunto de suposiciones ("categorías"), son ambas incongruentes, si se plantean así. Pues ellas mismas se apoyan en la verdaderamente colosal suposición de que es posible empezar a razonar sin suposiciones, o sólo con unas cuantas, y no obstante obtener resultados que valen la pena. (En realidad, este principio de soslayar toda presuposición no es, como algunos podrían pensar, un consejo para llegar a la perfección, sino una forma de la paradoja del mentiroso.)⁴

B dice: "Lo que dice A es falso".

Aplicando el método que hemos descrito más arriba, nos convencemos fácilmente de que cada uno de estos enunciados es paradójico. Consideremos ahora las siguientes dos frases, de las cuales la primera es el principio de que deberían gobernar los más sabios:

(A) El principio dice: Lo que los más sabios dicen según (B) debería ser la ley.

(B) Los más sabios dicen: Lo que el principio dice según (A) no debería ser la ley.

⁴ 1) Que el principio de evitar todas las presuposiciones es "una forma de la paradoja del mentiroso" en el sentido de la nota anterior, y, por tanto, contradictorio de sí mismo, se entenderá fácilmente si lo describimos así: Un filósofo empieza su investigación aceptando sin argumentaciones el principio de "Todos los principios aceptados sin argumentaciones son impermisibles". Está claro que, si aceptamos este principio como verdadero, debemos concluir, considerando lo que dice, que es impermissible. (La suposición contraria no plantea ninguna dificultad.) La observación "un consejo de perfección" alude a la acostumbrada crítica de este principio que sostuvo, por ejemplo, Husserl. J. Laird (en *Recent Philosophy*, 1936, p. 121) escribe acerca de este principio que es "un elemento cardinal en la filosofía de Husserl. Su éxito puede ser más dudoso, por cuanto las presuposiciones tienen una manera de colarse en él". Hasta aquí, estoy completamente de acuerdo; pero no tanto con la siguiente observación: "...evitar todas las presuposiciones bien puede ser un consejo de perfección, impracticable en un mundo que no las advierte".

2) Podemos considerar aquí otros "principios" que son, en el sentido de la nota anterior, "formas de la paradoja del mentiroso" y, por tanto, contradictorios de sí mismos.

a) Desde el punto de vista de la filosofía social, el siguiente "principio de sociologismo" (y su análogo "principio de historismo") resulta interesante: "Ninguna afirmación es absolutamente verdadera, y todas las afirmaciones son inevitablemente relativas al habitat social (o histórico) de quienes las originan". Está claro que las consideraciones hechas más arriba se aplican sin alteración. Porque si aceptamos que tal principio es verdadero, se colige que no es verdadero, sino sólo "relativo al habitat social o histórico de su originador". Véase también la nota 53 al capítulo 24 de *La sociedad abierta y sus enemigos*, y el texto de ese capítulo. [El capítulo de esta obra capital de Popper se titula "La filosofía oracular y la rebelión contra la razón". En la nota 53, Popper critica el historismo y el historicismo, particularmente el de Toynbee: "...podemos ahora arriesgarnos a efectuar un diagnóstico historioanalítico, y decir que el historismo es un producto típico, si bien algo anticuado, de nuestro tiempo, o mejor dicho, del retraso típico de las ciencias sociales de nuestro tiempo..." (N. del T.)]

b) Podemos encontrar algunos ejemplos de esta clase en el *Tractatus* de Wittgenstein. Uno es la proposición de Wittgenstein (citada más *in extenso* en la nota 13 al escrito 6): "la totalidad de las proposiciones verdaderas es [...] la totalidad de las ciencias naturales". Como esta proposición no pertenece a las ciencias naturales (sino más bien a una meta-ciencia, es decir, a una teoría que habla acerca de la ciencia), resulta que asevera su propia no verdad, y que, por tanto, es contradictoria.

Ahora bien, todo esto es un tanto abstracto, pero puede reformularse en conexión con el problema del racionalismo de manera menos formal. La actitud racionalista se caracteriza por la importancia que atribuye al argumento (la discusión) y a la experiencia. Pero ni la argumentación lógica ni la experiencia pueden establecer la actitud racionalista; pues sólo quienes estén dispuestos a considerar la argumentación o la experiencia, y que por tanto ya han adoptado esta actitud, serán convencidos por la argumentación y por la experimentación. En otras palabras, si

Además, resulta claro que esta proposición viola el propio principio de Wittgenstein (*Tractatus*, 3.332): "Ninguna proposición puede decir nada acerca de sí misma..."

Pero aun este último principio, al que llamaré "W", resulta ser una forma de la paradoja del mentiroso, y afirma su propia no verdad. (Por tanto, no puede ser —como Wittgenstein creía— equivalente, o un resumen, o un sustituto de "toda la teoría de tipos", es decir, de la teoría de Russell, diseñada para evitar las paradojas mediante la división de expresiones que parecen proposiciones en tres clases: proposiciones verdaderas, proposiciones falsas y expresiones sin significado, o pseudoproposiciones.) Porque el principio de Wittgenstein puede reformularse como sigue:

(W*) Cada expresión (y especialmente una que parezca una proposición) que contenga una referencia a sí misma —que contenga ya sea su propio nombre o una variable individual que pueda caber en una clase a la que ella misma pertenece— no es una proposición (sino una pseudoproposición sin significado).

Supongamos ahora que W* es verdadera. Entonces, considerando el hecho de que es una expresión, y que se refiere a cada expresión, no puede ser una proposición, y por tanto es, *a fortiori*, no verdadera.

Por tanto, la suposición de que es verdadera resulta insostenible; W+ no puede ser verdadera. Pero esto no demuestra que deba ser falsa; porque ni la suposición de que es falsa ni la de que es una expresión que carece de significado (o de que es absurda) nos mete en dificultades inmediatas.

Wittgenstein quizás habría dicho que se dio cuenta de esto mismo cuando escribió (6.54 [véase la nota 17 al escrito 6]): "mis proposiciones son elucidatorias en esta forma: quien me entiende reconoce finalmente que no tienen sentido..."; en todo caso, podemos conjeturar que se inclinaría por describir W* como sin significado, y no como falso. Sin embargo, pienso que no es que carezca de significado, sino que, simplemente, este principio es falso. O, más precisamente, pienso que en cada lenguaje formalizado (por ejemplo, uno en el que puede expresarse uno de los impronunciabiles asertos de Gödel) que contiene medios para hablar de sus propias expresiones y en los que tenemos nombres de clases de expresiones tales como "proposiciones" y "no proposiciones", la *formalización de una afirmación que, como W*, asevera su propio significado, será autocontradictoria, y no será ni carente de significado ni auténticamente paradójica*; será una proposición con significado, simplemente porque asevera de cada expresión de cierta clase que no es una proposición (es decir, una fórmula no bien formada); y tal afirmación será verdadera o falsa, pero no carente de significado, simplemente porque ser (o no ser) una bien formada proposición es una propiedad de las expresiones. Por ejemplo, "todas las expresiones carecen de significado" será autocontradictoria, pero no auténticamente paradójica, y así será también la aseveración "la expresión x carece de significado", si sustituimos "x" con un nombre de esta expresión. Modificando una idea de J. N. Findlay, podemos escribir:

La expresión que se obtiene al sustituir la variable en la siguiente expresión: "La expresión que se obtiene al sustituir la variable en la siguiente expresión con x el nombre de la cita de esta expresión, no es una aseveración", el nombre citado de esta expresión no es una aseveración.

Y lo que acabamos de escribir se convierte en una afirmación autocontradictoria. (Si escribimos dos veces "es una afirmación falsa", en vez de "no es una afirmación", obtene-

cualquier argumentación o experiencia ha de ser eficaz, debe adoptarse primero una actitud racionalista y, por tanto, la actitud racionalista no puede basarse en el argumento o la experimentación. (Y esta consideración es muy independiente de si existen o no argumentos racionales convincentes que favorezcan la adopción de la actitud racionalista.) Hemos de concluir de esto que ningún argumento racional tendrá un efecto racional en un hombre que no desee adoptar una actitud racional. Por consiguiente, un racionalismo total es insostenible.

mos una paradoja del mentiroso; si escribimos "es una afirmación no demostrable", obtenemos una afirmación gódeliana en escritura de Findlay.)

Resumamos. Al contrario de la primera impresión, descubrimos que una teoría que implica su propia carencia de significado no es sin significado, sino falsa, ya que el predicado "carente de significado", en oposición a "falso", no da lugar a paradojas. Y, por tanto, la teoría de Wittgenstein no carece de significado, como él creía, sino que es simplemente falsa (o, más específicamente, autocontradictoria).

3) Algunos positivistas han argüido que una tripartición de las expresiones de un lenguaje en (i) aseveraciones verdaderas, (ii) aseveraciones falsas, y (iii) expresiones carentes de significado (o, mejor, expresiones diferentes de las afirmaciones bien formadas) es más o menos "natural" y que proporciona, por su carencia de significado, la eliminación de paradojas y, al mismo tiempo, la eliminación de sistemas metah'sicos. El siguiente ejemplo puede demostrar que esta tripartición no basta. [Véase también *Conjeturas y refutaciones*, capítulo 14.]

Al oficial del general en jefe del contraespionaje se le dan tres cajas, con las etiquetas (i) "Caja del General", (ii) "Caja del Enemigo" (para que sea accesible a los espías del enemigo), y (iii) "Basura", y se le ordena distribuir toda la información que llegue antes de las 12 horas en punto entre estas tres cajas, según esta información sea: (i) verdadera (ii) falsa, o (iii) sin significado.

Este oficial recibe durante cierto tiempo información que puede distribuir fácilmente (entre ella, afirmaciones verdaderas de la teoría de los números naturales, etc., y quizá afirmaciones de lógica, como L: "De un conjunto de afirmaciones verdaderas, no puede derivarse válidamente ninguna afirmación falsa"). El último mensaje, L, que llega en el último correo, un poco antes de las 12, lo inquieta un poco, pues M dice así: "Del conjunto de todas las afirmaciones colocadas, o que se colocarán en la caja etiquetada "Caja del General", la afirmación "0=1" no puede derivarse válidamente". Primero, el oficial en jefe del contraespionaje vacila y piensa si no será conveniente colocar M en la caja (ii). Pero como se da cuenta de que, si coloca M en (ii), dará al enemigo valiosa información verdadera, decide por fin colocar M en (i).

Pero esto resulta un grave error. Porque los lógicos que operan con la lógica simbólica (¿expertos en logística?) del estado mayor general, después de formalizar (y *aritmétizar*) el contenido de la Caja del General, descubren que obtienen un conjunto de afirmaciones en donde hay una afirmación acerca de su propia coherencia; y esto, según el segundo teorema de Gödel sobre lo que puede decidirse, lleva a una contradicción, de manera que "0=1" en realidad puede deducirse de la supuesta información verídica proporcionada al general.

La solución de esta dificultad consiste en reconocer el hecho de que la exigencia de tripartición no está justificada, por lo menos en cuanto a lenguajes ordinarios; y podemos advertir, basados en la teoría de la verdad de Tarski, que no bastará ningún número de cajas para resolver esta dificultad. Al mismo tiempo, descubrimos que "la carencia de significado" en el sentido de "no pertenecer a las fórmulas bien formuladas" no es de ninguna manera un indicio de "lenguaje absurdo", en el sentido de "palabras que en realidad no significan nada, aunque pretendan ser profundamente significantes"; pero haber revelado que la metafísica tenía este carácter era el principal argumento de los positivistas.

Pero esto significa que quien adopte la actitud racionalista lo hace así porque ya ha adoptado, consciente o inconscientemente, alguna proposición, o decisión, o creencia, o conducta; adopción que puede llamarse "irracional". Ya sea que esta adopción sea provisional o conduzca a un hábito establecido, podemos describirla como una irracional *fe en la razón*. Así, el racionalismo está necesariamente lejos de lo totalizador o contenido en sí mismo. Esto con frecuencia ha sido pasado por alto por algunos racionalistas, que así se han expuesto a ser golpeados en su propio campo y con sus propias armas, cuando un irracionalista se ha tomado la molestia de arremeterla contra ellos. Y en verdad no escapó a la atención de algunos enemigos del racionalismo que uno puede siempre rechazar argumentos, ya sea todos los argumentos o los de cierta clase; y que tal actitud puede ser sostenida sin volverse lógicamente incongruente. Esto los hizo darse cuenta de que el racionalista acrítico, que cree que el racionalismo se contiene a sí mismo y que puede establecerse mediante la argumentación, puede estar en el error. El irracionalismo es lógicamente superior al racionalismo acrítico.

Entonces, ¿por qué no adoptar el irracionalismo? Muchos que empezaron como racionalistas, pero que se desilusionaron al descubrir que un racionalismo demasiado totalizador se derrota a sí mismo, han capitulado, en efecto, ante el irracionalismo. (Esto es lo que le ha sucedido a Whitehead,⁵ si no me equivoco.) Pero tal acción de pánico es del todo injustificada. Si bien un racionalismo acrítico y totalizador es lógicamente insostenible, y aunque un irracionalismo totalizador es lógicamente sostenible, esto no es razón para que debamos adoptar este último. Porque existen otras actitudes defendibles, sobre todo, la del racionalismo crítico, que reconoce el hecho de que la actitud racionalista fundamental resulta (al menos provisoriamente) de un acto de fe, de fe en la razón. En consecuencia, la opción está abierta ante nosotros. Podemos elegir abrazar alguna forma del irracionalismo, incluso una forma radical o totalizadora. Pero también tenemos la libertad de elegir una forma crítica de racionalismo, una que admita francamente su origen en una decisión irracional (y que, en esa medida, admita cierta prioridad de irracionalismo).

⁵ Es patente que la dificultad relacionada con el problema de la inducción fue lo que hizo que Whitehead desechara el argumento explicado en *Process and Reality* [Proceso y realidad], 1929. [En relación con todo este escrito, y con la posibilidad de que exista un racionalismo crítico exhaustivo, remitimos al lector a las obras de W. W. Bartley III, especialmente a *The Retreat to Commitment* [La retirada al compromiso], 1962. Véase también *La sociedad abierta y sus enemigos*, tomo II, Apéndice i.]

La elección que se nos presenta no es simplemente un asunto intelectual, ni cuestión de gusto. Constituye una decisión moral.⁶ Porque la cuestión de si adoptamos alguna forma más o menos radical de irracionalismo, o de si adoptamos esa mínima concesión al irracionalismo al que he llamado "racionalismo crítico", afectará profundamente toda nuestra actitud hacia los demás seres humanos, y hacia los problemas de la vida social. El racionalismo está estrechamente vinculado a la creencia en la unidad de la humanidad. El irracionalismo, que no está obligado a seguir ninguna regla de congruencia, puede combinarse con cualquier clase de creencia, incluso en la fraternidad de todos los hombres; pero el hecho de que puede combinarse fácilmente con una creencia muy diferente, y especialmente el hecho de que se presta fácilmente a apoyar la romántica creencia en la existencia de un cuerpo social elegido, en la división de los hombres en líderes y dirigidos, en amos naturales y esclavos naturales, demuestra claramente que hay una decisión moral implícita en la elección entre el irracionalismo y el racionalismo crítico.

Como ya lo he dicho (en *La sociedad abierta y sus enemigos*, capítulo 5), y como repito ahora en mi análisis de la versión acrítica del racionalismo, los argumentos no pueden *determinar* tan fundamental decisión moral. Pero esto no implica que nuestra elección no pueda ser, en absoluto, auxiliada por cualquier clase de argumento. Al contrario, siempre que nos enfrentamos a una decisión moral de índole más abstracta es muy útil analizar con todo cuidado las consecuencias que probablemente resulten de las opciones que tengamos a la mano. Porque sólo si podemos visualizar estas consecuencias de manera concreta y práctica sabremos realmente sobre qué será nuestra decisión; si no, decidiremos a ciegas. Para ilustrar este punto, puedo citar un pasaje de *Saint Joan* [Santa Juana], de Shaw. El que habla es el capellán; él ha pedido reiteradamente la muerte de Juana; pero cuando la ve en la hoguera se derrumba su ánimo y dice: "Yo no quería hacerle ningún daño. Yo no sabía cómo sería esto [...] No sabía lo que estaba haciendo [...] Si lo hubiera sabido, la habría arrancado de las manos de ellos. No lo sabes: no lo has visto: ¡es tan fácil hablar cuando no sabes! Te enajenas con las palabras

⁶ Se trata de una decisión moral, y no simplemente de una "cuestión de gusto", puesto que no es un asunto privado, sino que afecta a otros hombres y a sus vidas. La decisión a la que nos enfrentamos es importantísima desde el punto de vista de que los "enterados", que se enfrentan a ella, actúan como fideicomisarios intelectuales de quienes no tienen que enfrentarse a esta decisión.

[...], pero cuando entiendes por fin de qué se trata; cuando ves las cosas que has hecho, cuando eso te ciega, te tapa la nariz y te desgarran el corazón, entonces..., entonces..., ¡oh Dios!, ¡aparta de mí este espectáculo!" En la pieza de Shaw hay, por supuesto, otros personajes que sabían exactamente lo que estaban haciendo y, no obstante, decidieron hacerlo, y después no lo lamentaron. A algunas personas les disgusta ver arder en la hoguera a sus prójimos, y a otras no. Este punto (que pasaron por alto muchos optimistas Victorianos) es importante, pues demuestra que el análisis racional de las consecuencias de una decisión no convierte en racional esa decisión; las consecuencias no determinan nuestra decisión; somos nosotros siempre quienes decidimos. Pero un análisis de las consecuencias concretas, y su clara toma de conciencia en lo que llamamos nuestra "imaginación", establece la diferencia entre tomar una decisión a ciegas y tomarla con los ojos abiertos: y como utilizamos muy poco la imaginación,⁷ con demasiada frecuencia decidimos a ciegas. Esto es así, sobre todo, si estamos intoxicados por una filosofía sentenciosa, uno de los medios más potentes para enajenarnos con palabras, para emplear la expresión de Shaw.

El análisis racional e imaginativo de las consecuencias de una teoría moral presenta cierta analogía con el método científico. Porque también en la ciencia sucede que no aceptamos una teoría abstracta porque sea convincente en sí misma; decidimos aceptarla o rechazarla después de haber investigado aquellas consecuencias concretas y prácticas que pueden probarse más exactamente mediante la experimentación. Pero existe una diferencia fundamental a este respecto: en el caso de la teoría científica, nuestra decisión depende de los resultados de los experimentos. Si éstos confirman la teoría, podemos aceptarla hasta que descubramos otra mejor. Si la contradicen, la rechazamos. Sin embargo, en el caso de una teoría moral, sólo podemos confrontar sus consecuencias con nuestra conciencia. Y aunque el veredicto de la experimentación no depende de nosotros, el veredicto de nuestra conciencia sí.

Espero haber esclarecido en qué sentido puede influir el análisis de las consecuencias en nuestra decisión sin determinarla. Y al presentar las consecuencias de las dos opciones entre las que debemos decidir, el racionalismo y el irracionalismo, advierto al lector que en esto seré parcial. Hasta este punto, al presentar las dos opciones de la decisión moral que se presenta ante nosotros —es, en muchos aspectos, la decisión más fundamental en el campo de la ética—, he tratado de ser imparcial, aunque no he ocultado hacia dónde se inclinan mis simpatías. Pero

⁷ En mi opinión, la mayor fuerza del cristianismo quizá sea que aparezca fundamentalmente no como especulación abstracta, sino ante la imaginación, al describir de manera muy concreta el sufrimiento del hombre.

ahora presentaré aquellas consideraciones de las consecuencias de las dos opciones que me parecen más reveladoras y que han influido en mí para rechazar el irracionalismo y aceptar la fe en la razón.

Examinemos primero las consecuencias del irracionalismo. El irracionalista insiste en que las emociones y las pasiones, y no la razón, constituyen los resortes principales de la acción humana. A la respuesta del racionalista de que, aunque esto sea así, deberíamos hacer lo que podamos para remediarlo, y procurar que la razón desempeñe un papel tan importante como sea posible, el irracionalista contrarreplica (si condesciende a entablar una discusión) que esta actitud es irremediablemente poco realista. Porque no toma en cuenta la debilidad de la "naturaleza humana", la débil dotación intelectual de la mayoría de los hombres, su obvia dependencia de las emociones y pasiones.

Es mi firme convicción que el acento que ponen los irracionistas en la emoción y en la pasión nos conduce en última instancia a lo que sólo puedo calificar de crimen. Una razón de esta opinión es que tal actitud, que en el mejor de los casos representa resignarse ante la naturaleza irracional de los seres humanos, y en el peor, el desprecio a la razón humana, debe tener que apelar a la violencia y a la fuerza bruta como arbitro supremo en las disputas. Porque, si surge una disputa, ello significará que aquellas emociones y pasiones más constructivas que en principio ayudarían a dirimirla: la reverencia, el amor, la devoción hacia una causa común, etc., habrán sido incapaces de resolver el problema. Y si esto es así, entonces, ¿qué le queda al irracionalista, excepto recurrir a otras emociones y pasiones menos constructivas, como son el temor, el odio, la envidia y, al final, la violencia? Esta tendencia se ve muy fortalecida por otra actitud, quizá más importante, que en mi opinión también es inherente al irracionalismo; a saber, la de acentuar la desigualdad de los hombres.

Por supuesto, es innegable que los individuos humanos, como todas las demás entidades de nuestro mundo, son en muchos aspectos muy desiguales. Tampoco es posible dudar de que esta desigualdad tiene gran importancia y de que, en muchos aspectos, resulta muy deseable.⁸ (El temor a que el desarrollo de la producción en masa y la colectivización reaccionen sobre el hombre destruyendo su desigualdad o su individualidad es una de las pesadillas⁹ de nuestro tiempo.) Pero todo esto simplemente no tiene peso en la cuestión de si deberíamos decidir

⁸ Kant, el gran igualitario en materia de decisiones morales, hizo hincapié en las bendiciones implícitas en el hecho de la desigualdad humana. Vio en la variedad y en la individualidad de los caracteres y en las opiniones de los humanos una de las principales condiciones del progreso, tanto moral como material.

⁹ La alusión es a *Brave New World* [*Un mundo feliz*], novela de Aldous Huxley, 1932.

tratar a los hombres, sobre todo en lo político, como a iguales o con tanta igualdad como sea posible; es decir, como a seres que poseen iguales derechos e igual exigencia a que se les trate en forma igualitaria; y tampoco debe incidir esta comprobación de la desigualdad en la cuestión de si deberíamos establecer instituciones políticas basadas en la igualdad. "La igualdad ante la ley" *no es un hecho, sino una exigencia política*¹⁰ basada en una decisión moral; y es muy independiente de la teoría —probablemente falsa— de que "todos los hombres nacen iguales en derechos". Ahora bien, no pretendo decir que la adopción de esta actitud humanitaria de imparcialidad sea la consecuencia directa de una decisión en favor del racionalismo. Pero la tendencia a la imparcialidad está estrechamente relacionada con el racionalismo, y difícilmente puede excluirse del credo racionalista. Tampoco intento decir que un irracionalista no podría adoptar consistentemente una actitud igualitaria o imparcial; y aunque no pudiera hacerlo consistentemente, el irracionalista no está obligado a ser consistente. Pero deseo hacer hincapié en que la actitud irracionalista no puede evitar, casi, caer en la actitud opuesta al igualitarismo. Esto se relaciona con el acento que pone en las emociones y en las pasiones; porque no podemos sentir las mismas emociones hacia todo el mundo. Emocionalmente, dividimos a los hombres en los que están cerca de nosotros y los que están lejos de nosotros. La división de la humanidad entre amigos y enemigos es casi una división emocional obvia: y esta división está incluso reconocida en el mandamiento cristiano de "¡Ama a tus enemigos!" Hasta el mejor cristiano que realmente vive según este mandamiento (no hay muchos, a juzgar por la actitud que tiene el buen cristiano promedio hacia los "materialistas" y los "ateos") no puede sentir igual amor por todos los hombres. No podemos amar realmente "en lo abstracto"; sólo podemos amar a quienes conocemos. Así pues, recurrir incluso a nuestras mejores emociones, el amor y la compasión, sólo puede tender a dividir a la humanidad en diferentes categorías. Y esto puede ser más cierto si apelamos a emociones y pasiones inferiores. Nuestra reacción "natural" será dividir a la humanidad en amigos y enemigos; entre aquellos que pertenecen a nuestra tribu, a nuestra comunidad emocional, y aquellos que están fuera de ella; entre creyentes y no creyentes, entre compatrio-

¹⁰ Para la distinción entre hechos, y decisiones o exigencias, véase el capítulo 5 de *La sociedad abierta y sus enemigos*. En cuanto al "lenguaje de las exigencias políticas" (o "propuestas" en el sentido que les da L. J. Russell), véase la *op. cit.*, capítulo 6, sección vi.

Me inclino por afirmar que la teoría de la innata igualdad intelectual de todos los hombres es falsa; pero dado que hombres como Niels Bohr aseveran que la influencia del medio es la única responsable de las diferencias individuales, y como no hay suficientes datos experimentales para decidir acerca de este punto, quizás habría sido mejor decir "probablemente falsa".

tas y extranjeros, entre camaradas de la misma clase y enemigos de clase, entre líderes y conducidos.

Ya he mencionado que la teoría de que nuestros pensamientos y opiniones dependen de nuestra situación en determinada clase social o de nuestros intereses nacionales debe conducirnos hacia el irracionalismo. Ahora deseo hacer hincapié en que lo contrario también es verdad. El abandono de la actitud racionalista, del respeto por la razón y por la argumentación y los puntos de vista del interlocutor, el acento puesto en las "más profundas" capas de la naturaleza humana, todo esto debe conducirnos al punto de vista de que el pensamiento es sólo la manifestación un tanto superficial de lo que reside en estas irracionales profundidades. Creo que este enfoque casi siempre origina una actitud que considera a la persona del que piensa, en vez de considerar su pensamiento. Debe originar la creencia de que "pensamos con nuestra sangre", o con "nuestro legado nacional", o con "nuestra clase social". Este punto de vista puede presentarse en forma materialista o de manera muy espiritual; la idea de que "pensamos con nuestra raza" acaso sea sustituida por la idea de que hay almas elegidas o inspiradas que "piensan por la gracia de Dios". Me niego, basado en principios morales, a ser convencido por estas diferencias; porque la decisiva similitud entre todos estos intelectualmente inmodestos puntos de vista es que no juzgan un pensamiento por sus propios méritos. Al abandonar así la razón, dividen a la humanidad en amigos y enemigos; entre los muy pocos que comparten la razón con los dioses, y los muchos que no (como dice Platón); entre los muy pocos que están cerca y los muchos que están lejos; entre los que hablan la intraducible lengua de nuestras propias emociones y pasiones y aquellos cuya lengua no es nuestra lengua. En cuanto hemos caído en esta actitud, el igualitarismo político se vuelve casi imposible.

Ahora bien, la adopción de una actitud antiigualitaria en la vida política, es decir, en el campo de los problemas que se refieren al poder del hombre sobre el hombre, es precisamente lo que yo calificaría de criminal. Porque da una justificación a la actitud que considera que diferentes categorías de personas tienen diferentes derechos; que el amo tiene derecho a esclavizar al esclavo, que algunos hombres tienen derecho a usar a otros como si fueran herramienta. Y en última instancia, esta actitud antiigualitaria se utilizará, como lo hizo Platón,¹¹ para justificar el asesinato.

¹¹ Véase, por ejemplo, *Statesman [El estadista]*, 293 C-E. Otro pasaje de éstos es *La República*, 409 E-410 A. Después de hablar (en 409 B/C) del "buen juez [...] que es bueno por la bondad de su alma", Platón prosigue: "¿Y no vais a instaurar médicos y jueces [...] que cuidarán de aquellos ciudadanos cuya constitución física y mental es saludable y

No paso por alto el hecho de que existen irracionalistas que aman a la humanidad, ni que no todas las formas de irracionalismo engendran criminalidad. Pero sostengo que quien enseña que no es la razón, sino el amor, el que debe gobernar abre el camino para quienes gobiernan mediante el odio. (Creo que fue Sócrates el que se dio cuenta de esto cuando sugirió¹² que la desconfianza o el odio a la argumentación se relaciona con la desconfianza o el odio hacia los hombres.) Quienes no ven de inmediato esta conexión, quienes creen en un gobierno directo del amor emocional, deberían considerar que tal amor, en sí mismo, en verdad no promueve la imparcialidad. Y tampoco puede dirimir los conflictos. Que el amor como tal puede ser incapaz de arreglar un conflicto se demuestra al ponderar un inocuo caso de prueba, que puede considerarse representativo de otros casos más graves. A Tomás le gusta el teatro, y Daniel es muy aficionado al baile. Tomás insiste amablemente en ir a bailar, mientras que Daniel insiste en que vayan al teatro para agradar a Tomás. Este conflicto de intereses no puede dirimirse mediante el cariño o el amor; más bien, cuanto más grande sea el amor, más fuerte será el conflicto. Hay sólo dos soluciones posibles: el empleo de la emoción, y en último término de la violencia, y el empleo de la razón, de la imparcialidad, del compromiso o la componenda razonable. Todo esto no tiene la intención de indicar que no aprecio la diferencia entre el amor y el odio, o de que crea que la vida merezca la pena de vivirse sin amor. (Y estoy dispuesto a aceptar que la idea cristiana del amor no tiene sólo un significado emocional.) Pero insisto en que ninguna emoción, ni siquiera el amor, puede sustituir el gobierno de instituciones controladas por la razón.

Éste, por supuesto, no es el único argumento en contra de la idea de la regla moral del amor. Amar a una persona significa desear hacerla feliz. (De paso, recordemos que ésta fue la definición del amor que dio Tomás de Aquino.) Pero, de todos los ideales políticos, el de hacer feliz al pueblo quizá sea el más peligroso. Conduce invariablemente a intentar imponer a los demás nuestra escala de "más altos" valores para hacer que realicen lo que a nosotros nos parece de la mayor importancia para su felicidad; como si dijéramos, para salvar sus almas. Y esto nos lleva hacia lo utópico y hacia el romanticismo. Todos estamos convencidos de que todo el mundo sería feliz en la hermosa y perfecta comunidad de nuestros sueños. Y no cabe la menor duda de que habría un cielo en la tierra si todos fuéramos capaces de amarnos los unos a

buena? Aquellos cuya salud física sea mala, se dejarán morir. Y aquellos cuya alma es de mala naturaleza e incurable, realmente matarán". —"Sí —dijo—, puesto que has probado que esto es lo mejor, tanto para aquellos a quienes sucede, como para el Estado".

¹² Véase *Fedón*, 89 C/D.

los otros. Pero el intento de hacer que haya cielo en la tierra produce invariablemente un infierno [véase más adelante el escrito 24]. Conduce a la intolerancia. Desemboca en guerras de religión, y en la salvación de las almas por la Inquisición. Porque, estoy convencido de ello, esta idea se basa en una total falta de comprensión de nuestros deberes morales. Es nuestro deber ayudar a quienes necesitan de nuestra ayuda; pero no puede ser nuestro deber hacer felices a los demás, puesto que esto no depende de nosotros, y puesto que, las más de las veces, esto significaría inmiscuirnos en la vida privada de aquellos para los que tenemos tan amables intenciones. La exigencia política de métodos poco sistemáticos (en oposición a los métodos utópicos) corresponde a la decisión de que hemos de considerar un deber luchar contra el sufrimiento, mientras que el derecho a velar por la felicidad de los demás ha de considerarse un privilegio confinado al estrecho círculo de los amigos. En el caso de estos amigos, quizá tengamos cierto derecho a intentar imponer nuestra escala de valores; nuestras preferencias en cuanto a escuchar cierta clase de música, por ejemplo. (Y acaso sintamos que es nuestro deber abrir para ellos un mundo de valores que, así esperamos, pueda contribuir en tan gran medida a hacerlos felices.) Este derecho nuestro existe sólo a condición de que nuestros amigos puedan en todo momento prescindir de nosotros; porque las amistades pueden terminar. Pero la utilización de medios políticos para imponer a los demás nuestra escala de valores es otra cosa muy diferente. El dolor, el sufrimiento, la injusticia, y la prevención de estas calamidades, tales son los eternos problemas de la moral pública, la "agenda" de la política de gobierno (como habría dicho Bentham). Los valores "más altos" deberían considerarse en gran medida como "fuera de la agenda", y deberían dejarse en el reino del *laissez-faire*. Así, deberíamos decir: Ayuda a tus enemigos, asiste a los que sufren, incluso si te odian; pero ama sólo a tus amigos.

Ésta es únicamente una parte del proceso contra el irracionalismo, y contra las consecuencias que me inducen a adoptar la actitud contraria; es decir, el racionalismo crítico. Esta última actitud, con su hincapié en la argumentación y la experimentación, con su lema: "Es posible que yo esté en el error, y acaso tú estés en lo correcto, en lo cierto; y si ambos hacemos un esfuerzo, quizá podamos acercarnos más a la verdad", es, como ya he dicho, muy afín a la actitud científica. Está muy ligada a la idea de que todo el mundo está expuesto a cometer errores, los cuales puede detectar la persona misma que los comete, u otras personas, o la persona misma con la ayuda de la crítica de otras personas. Por tanto, propone la idea de que nadie debería ser su propio juez, y también sugiere la idea de la imparcialidad. (Esto se relaciona estrechamente con la

idea de la "objetividad científica" [analizada más adelante, en el escrito 30].) Su fe en la razón es una fe, no sólo en nuestra propia razón, sino también —e incluso en mayor medida— en la de otros. Así, un racionalista, aun si se cree intelectualmente superior a otros, rechazará toda pretensión de ser una autoridad,¹³ pues sabe perfectamente que, si su inteligencia es superior a la de otros (lo cual sería para él muy difícil de juzgar), lo es únicamente en tanto sea capaz de aprender de la crítica, así como de los errores propios y de los ajenos, y que sólo podemos aprender en esta forma si tomamos en serio a los otros y sus argumentos. Por consiguiente, el racionalismo está imbuido de la idea de que el otro ser humano tiene derecho a que lo escuchen y a defender sus argumentos. Implica, por ello, el reconocimiento de la tolerancia, por lo menos¹⁴ de quienes no son ellos mismos intolerantes. Nadie mata a un hombre cuando se adopta la actitud de, antes de actuar, escuchar sus argumentos. (Kant estaba en lo correcto cuando basó su "Regla de Oro" en la idea de la razón. No cabe duda de que es imposible probar la rectitud de cualquier principio ético, o siquiera argumentar en su favor tal como argüimos en favor de una aseveración científica. La ética no es una ciencia. Pero aunque no existe una "base científica racional" de la ética, sí existe una base ética de la ciencia y del racionalismo.) La idea de imparcialidad conduce también a la de responsabilidad; no sólo tenemos que escuchar los argumentos, sino que también tenemos el deber de responder, de contestar, cuando nuestras acciones afectan a otros seres humanos. En última instancia, de esta manera el racionalismo está ligado al reconocimiento de la necesidad de instaurar instituciones sociales para proteger la libertad de crítica, la libertad de pensamiento y, por ende, la libertad de los hombres. Y establece una especie de obligación moral tendiente a apoyar estas instituciones. Por ello, el racionalismo está estrechamente vinculado a la exigencia política de emprender una ingeniería social práctica —ingeniería de paso a paso, por supuesto— en el sentido humanitario, con la exigencia de racionalización de la sociedad,¹⁵ con la planificación para la libertad, y que tienda al control por medio de la razón, no por medio de la "ciencia" ni por medio de

¹³ Un ejemplo de esto es H. G. Wells, que puso al primer capítulo de su libro *The Common Sense of War and Peace* [*El sentido común de la guerra y la paz*], 1940, este excelente título: "Los hombres hechos y derechos no necesitan líderes".

¹⁴ En cuanto al problema y la paradoja de la tolerancia [véase la nota 4 al escrito 25].

¹⁵ El "mundo" no es racional, pero es tarea de la ciencia racionalizarlo. La "sociedad" no es racional, pero es tarea del ingeniero social racionalizarla. (Esto no significa, por supuesto, que deberíamos "dirigirla", o que sea deseable la "planificación" centralizada o colectivista.) El lenguaje ordinario no es racional, pero es nuestra tarea racionalizarlo o, por lo menos, conservar sus normas de claridad. La actitud que aquí hemos caracterizado podría describirse como "racionalismo pragmático". Este racionalismo pragmático tiene que ver con un racionalismo no crítico y con el irracionalismo de manera semejante a como se

una autoridad platónica y seudorracional, sino por medio de esa razón socrática que está consciente de sus limitaciones, y que por tanto respeta al otro nombre y no aspira a ejercer coerción en él, ni siquiera para procurar su felicidad. La adopción del racionalismo implica, además, que existe un medio de comunicación común, un lenguaje de razón común; establece una especie de obligación moral respecto a ese lenguaje, la obligación de conservar sus normas de claridad¹⁶ y de emplearlo de tal modo que retenga su función de vehículo de la argumentación. Es decir, la obligación de utilizarlo en forma clara; de utilizarlo como instrumento de comunicación racional, o de información significativa, y no como un medio de "autoexpresión", como la viciosa jerga romántica de la mayoría de nuestros "educacionalistas". (Es característico de la moderna histeria romántica combinar el colectivismo hegeliano referente a la "razón" con un excesivo individualismo concerniente a las "emociones": así, se pone énfasis en el lenguaje como medio de autoexpresión, en vez de como un medio de comunicación. Ambas actitudes, por supuesto, son parte de la revuelta en contra de la razón.) Ello implica el reconocimiento de que la humanidad está unida por el hecho de que nuestras distintas lenguas maternas, en la medida en que son racionales, pueden traducirse unas a otras. Se reconoce así la unidad de la razón humana.

Podemos añadir algunas observaciones sobre la relación de la actitud racionalista con la actitud de estar dispuesto a utilizar lo que generalmente se llama "imaginación". Se supone con frecuencia que la imaginación tiene una estrecha afinidad con la emoción y, por tanto, con el irracionalismo, y que el racionalismo tiende más bien a ser un escolasticismo seco y sin imaginación. No sé si tal suposición tenga alguna base psicológica, y dudo que así sea. Pero mis intereses son institucionales antes que psicológicos, y desde el punto de vista institucional (y también desde el enfoque metodológico) es obvio que el racionalismo debe alentar el empleo de la imaginación, porque la necesita, en tanto que el

relaciona el racionalismo crítico con estos dos últimos. Porque el racionalismo no crítico puede argumentar que el mundo es racional y que la tarea de la ciencia consiste en descubrir su racionalidad, mientras que el irracionalista insistirá, acaso, en que siendo el mundo fundamentalmente irracional, debería experimentarse y explicarse plenamente por medio de emociones y pasiones (o mediante nuestra intuición intelectual), más que por medio de métodos científicos. Opuesto a esto, el racionalismo pragmático puede reconocer que el mundo no es racional, pero exige que lo sometamos a la razón, tanto como sea posible. En palabras de Carnap (*The Logical Structure of the World* [*La estructura lógica del mundo*], 1928, p. vi; traducción inglesa, 1967, p. xviii), podríamos describir lo que llamo "racionalismo pragmático" como "la actitud que ansia claridad en todo, pero que reconoce lo nunca plenamente comprensible o lo nunca plenamente racional de lo intrincado de los acontecimientos de la vida".

¹⁶ [Véase también la nota 16 al escrito 6.]

irracionalismo tiende a desalentar lo imaginativo. El hecho mismo de que el racionalismo es crítico, mientras que el irracionalismo debe tender hacia el dogmatismo (donde no hay argumentación posible, y no queda sino la aceptación plena o la negación rotunda), nos orienta al respecto. La crítica pide siempre cierto grado de imaginación, mientras que el dogmatismo la suprime. De manera semejante, la investigación científica y la tecnología de construcción e invención son inconcebibles sin un empleo considerable de la imaginación; uno debe ofrecer algo nuevo en estos campos (opuesto al campo de la filosofía "sentenciosa", donde una repetición sin fin de palabras impresionantes parece resolver el problema). De igual importancia, por lo menos, es la parte que desempeña la imaginación en la aplicación práctica del igualitarismo y de la imparcialidad. La actitud básica del racionalista, "Es posible que yo esté en el error, y que tú estés en lo cierto", exige, al ponerse en práctica, y especialmente cuando están de por medio conflictos humanos, un verdadero esfuerzo de nuestra imaginación. Admito que las emociones de amor y compasión a veces pueden suscitar un esfuerzo similar al de la imaginación. Pero sostengo que es humanamente imposible amar a un gran número de personas o sufrir con ellas; ni tampoco me parece deseable que lo hiciéramos, ya que a la postre esto destruiría nuestra capacidad para ayudar o la intensidad de estas mismas emociones. En cambio, la razón, apoyada por la imaginación, nos permite entender que los hombres que están muy lejos de nosotros, a los que nunca veremos, son como nosotros, y que sus interrelaciones son como las nuestras con aquellos a quienes amamos. Una actitud emocional directa hacia el conjunto abstracto de toda la humanidad me parece casi imposible. Podemos amar a la humanidad sólo en ciertos individuos concretos. Pero, utilizando el pensamiento y la imaginación, podemos estar dispuestos a ayudar a quienes necesiten nuestra ayuda.

Todas estas consideraciones demuestran, en mi opinión, que el nexo entre el racionalismo y el humanitarismo es muy fuerte, y ciertamente mucho más fuerte que el correspondiente enmarañamiento del irracionalismo con el antiigualitarismo y la actitud antihumanitaria. Estoy convencido de que, en la medida de lo posible, esta consecuencia es corroborada por la experiencia. La actitud racionalista parece combinarse casi siempre con un concepto de la vida básicamente igualitario y humanitario; por otra parte, el irracionalismo despliega en la mayoría de los casos por lo menos alguna de las tendencias antiigualitarias descritas, aunque también esté a menudo asociado con el humanitarismo. Estoy convencido de que esta última conexión se fundamenta bien en los hechos.

3. CONOCIMIENTO SIN AUTORIDAD (1960)

ESTA parte de mi conferencia podría describirse como un ataque al *empirismo*, como lo formula, por ejemplo, la siguiente aseveración clásica de Hume:

Si pregunto por qué cree usted en algún hecho particular..., debe usted darme alguna razón de ello: y esta razón será algún otro hecho, relacionado con él. Pero como no puede usted proseguir de esta manera, *in infinitum*, debe usted concluir por lo menos con un hecho, que esté presente en su memoria o en sus sentidos; o debe usted reconocer que su creencia no tiene ningún fundamento.¹

El problema de la validez del empirismo puede expresarse, en términos generales, como sigue: ¿es la observación la fuente fundamental de nuestro conocimiento de la naturaleza? Y de no ser así, ¿cuáles son las fuentes de nuestro conocimiento?

Estas preguntas siguen en pie, no obstante cuanto haya yo dicho acerca de Bacon, y aunque haya convertido aquellas partes de su filosofía que he comentado en algo poco atractivo para los baconianos y para otros cultivadores de la filosofía empírica.

El problema de la fuente de nuestro conocimiento se ha replanteado recientemente así: Si hacemos una afirmación, debemos justificarla; y esto significa que debemos contestar las preguntas siguientes:

¿Cómo lo sabe? ¿Cuáles son las fuentes de su afirmación? Esto, según sostiene el empírico, equivale a su vez a otra pregunta: ¿Qué observaciones (o recuerdos de observaciones) subyacen en su afirmación? Considero que esta sarta de preguntas es un tanto insatisfactoria.

Ante todo, la mayoría de nuestras afirmaciones no se basan en observaciones, sino en toda clase de otras fuentes de conocimiento. "Lo leí en *The Times*", o quizá "Lo leí en la *Encyclopaedia Britannica*", es una respuesta más plausible o más definitiva a la pregunta de "¿Cómo lo sabe

¹ Véase D. Hume, *An Enquiry Concerning Human Understanding* [Ensayos sobre el entendimiento humano], 1748, sección v, Primera Parte; edición de L. Selby-Bigge, p. 46. Véase también la aseveración de Hume de que "es imposible para nosotros pensar en nada que no hayamos antes sentido, ya sea mediante nuestros sentidos externos o internos", en la sección vil, Primera Parte; edición de Selby-Bigge, p. 62.

usted?", que "Lo he observado" o "Lo sé por una observación que hice el año pasado".

El empírico replicará: "Pero, ¿cómo piensa usted que *The Times* o la *Encyclopaedia Britannica* obtuvieron su información? Seguramente, si lleva a cabo su encuesta durante un buen tiempo, terminará usted por llegar a *informes de las observaciones de testigos oculares* (a veces llamados 'frases de protocolo' o —si usted mismo investigó— 'afirmaciones básicas')". El empírico proseguirá: "Cierto es que los libros están hechos en gran medida a partir de otros libros. También es cierto que un historiador, por ejemplo, trabajará a partir de documentos. Pero en definitiva, en último análisis, estos otros libros, o estos documentos, deben de haberse basado en observaciones. De otra manera, tendrían que describirse como poesía, o invenciones, o mentiras, y no como testimonios. En este sentido, nosotros, los empíricos, aseveramos que la observación debe ser la fuente última de nuestro conocimiento".

Aquí tenemos el argumento de los empíricos, tal como aún lo presentan algunos de mis amigos positivistas.

Trataré de demostrar que este argumento es tan poco válido como el de Bacon; que la respuesta a la pregunta acerca de las fuentes del conocimiento va en contra del empírico; y, por último, que todo este asunto de las fuentes fundamentales —fuentes a las que puede uno apelar, como podría hacerlo ante un alto tribunal o ante una autoridad superior— debe rechazarse, por estar basado en una equivocación.

Primero deseo demostrar que si usted, en efecto, sigue preguntando a *The Times* y a sus corresponsales acerca de las fuentes de su conocimiento, en realidad jamás llegará, en absoluto, a aquellas observaciones por testigos oculares, en cuya existencia cree el empírico. Más bien descubrirá usted que, a cada paso que dé en este sentido, aumentará la necesidad de dar más pasos, como si hubiera lanzado a rodar una bola de nieve.

Tomemos como ejemplo la clase de afirmación por la cual las personas razonables aceptarían simplemente como satisfactoria la respuesta "Lo leí en *The Times*"; digamos que se trata de "El Primer Ministro ha decidido regresar a Londres varios días antes de lo programado". Ahora, supongamos por un momento que alguien duda de esta noticia, o siente la necesidad de investigar más a fondo su veracidad. ¿Qué hará? Si tiene un amigo en la oficina del Primer Ministro, el método más sencillo y directo de averiguarlo sería llamarle por teléfono; y si este amigo suyo corrobora el mensaje, se dará por satisfecho.

En otras palabras, el investigador intentará, si le es posible hacerlo, comprobar o examinar *el hecho mismo a que se refiere la afirmación*, más que seguirle el rastro a la fuente de información. Pero según la teoría empírica, la aseveración "Lo he leído en *The Times*" es sólo un primer

paso en el procedimiento de justificación que consiste en seguir el rastro de la fuente de información esencial. ¿Cuál es el siguiente paso?

Hay por lo menos dos pasos siguientes. Uno sería pensar que "Lo he leído en *The Times*" es a su vez una afirmación, y que podríamos preguntar: "¿Cuál es la fuente de su conocimiento de que leyó eso en *The Times*, y no, digamos, en otro periódico de apariencia muy similar a *The Times*?" El otro paso consiste en preguntar al periódico *The Times* las fuentes de su conocimiento. La respuesta a la primera pregunta podría ser: "Pues sucede que nosotros sólo estamos suscritos a *The Times*, y lo recibimos siempre por la mañana", lo cual da lugar a otras muchas preguntas acerca de las fuentes de información, cuyo rastro no seguiremos. La segunda pregunta puede dar pie a que el director de *The Times* responda: "Recibimos un telefonema de la oficina del Primer Ministro". Ahora bien, según el procedimiento del empírico, en esta etapa deberíamos preguntar: "¿Quién es el caballero que atendió la llamada telefónica?", y luego obtener su informe por escrito; pero también tendríamos que preguntar a ese caballero: "¿Cuál es la fuente de su conocimiento de que la voz que oyó usted procedía efectivamente de un funcionario de la oficina del Primer Ministro?", etcétera.

Existe una razón sencilla por la que esta tediosa secuencia de preguntas jamás llega a una conclusión satisfactoria; a saber: cada testigo tiene que hacer siempre un amplio uso, en su informe, de su conocimiento de personas, lugares, cosas, maneras de utilizar el lenguaje, convencionalismos sociales, etc. No puede confiar simplemente en sus ojos o en sus oídos, especialmente si su informe se usará para justificar cualquier aseveración que necesite justificarse. Pero este hecho, por supuesto, suscitará siempre nuevas preguntas en cuanto a las fuentes de aquellos elementos de su conocimiento que no son de inmediata observación.

Por todo esto, el programa de seguir el rastro de todo conocimiento hasta sus orígenes, en la observación, es lógicamente imposible de llevar a cabo en toda su magnitud: nos llevaría hasta una regresión infinita. (La doctrina de que la verdad es manifiesta interrumpe la regresión. Esto es interesante, porque puede ayudar a explicar el atractivo que ejerce esa doctrina.)

Deseo mencionar, entre paréntesis, que este argumento está íntimamente relacionado con otro: que toda observación implica una interpretación, a la luz de nuestro conocimiento teórico, o de que el conocimiento puramente observacional, no adulterado por la teoría, en caso de que esto fuera posible, sería del todo estéril y fútil. [Véase, más adelante, el último párrafo del escrito 11, sección i.]²

² Véase también *La lógica de la investigación científica*, Apéndice *x (2).

Lo más sobresaliente acerca del programa "observacionista" de preguntar por las fuentes de información —aparte del tedio que causa— es su flagrante violación del sentido común. Porque si dudamos de la veracidad de una afirmación, el procedimiento normal sería ponerla a prueba, y no preguntar por sus fuentes de información; y si descubrimos alguna corroboración independiente, aceptaremos a menudo esa aseveración sin más, sin molestarnos en absoluto por averiguar sus fuentes.

Por supuesto, en algunos casos la situación es diferente. Comprobar una afirmación *histórica* significa siempre remontarse a sus fuentes; pero, como regla general, no recurriremos a los informes de los testigos oculares.

Claro está que ningún historiador aceptará la autenticidad de documentos sin llevar a cabo un examen crítico de ellos. Hay problemas de autenticidad y problemas de tendencia partidaria, y también existen problemas como los de reconstruir una fuente de información de hace mucho tiempo. Y hay además otras dificultades; por ejemplo: ¿estaba presente el escritor cuando ocurrieron estos acontecimientos? Sin embargo, éste no es uno de los problemas característicos del historiador. Acaso le preocupe la confiabilidad de un informe, pero rara vez le preocupará si el redactor de un documento era un testigo ocular del suceso en cuestión, aun suponiendo que este acontecimiento fuera observable. Una carta que diga "Ayer cambié de opinión acerca de este asunto" puede ser la más valiosa prueba histórica, si bien los cambios de opinión no son sucesos observables (y aunque podamos conjeturar, a la vista de otras pruebas, que el redactor de esa carta estaba mintiendo).

En cuanto a los testigos oculares, son importantes casi exclusivamente en el tribunal, donde se les puede someter a careos. Como bien saben los abogados, los testigos oculares suelen equivocarse. Esto se ha investigado experimentalmente, con los más asombrosos resultados. Los testigos más deseosos de describir un acontecimiento tal como ocurrió están sujetos a cometer muchísimos errores, especialmente si ocurrió algo emocionante en un abrir y cerrar de ojos; y si un suceso induce a hacer alguna interpretación atractiva, esta interpretación, las más de las veces, es una distorsión de lo que en realidad se presenció.

El concepto de Hume sobre el conocimiento en historia era diferente:

...Creemos —escribe en su *Tratado*— que César fue asesinado en el recinto del Senado, en los idus de marzo... porque este hecho está comprobado por el testimonio unánime de historiadores, quienes concuerdan en asignar este preciso tiempo y lugar a ese acontecimiento. Aquí tenemos ciertos signos y ciertas cartas presentes, ya sea en nuestra memoria o en nuestros sentidos; caracteres escritos que también recordamos haber sido utilizados como signos de ciertas ideas: y estas ideas estuvieron, o en las mentes de quienes estu-

vieron inmediatamente presentes en ese acto, y recibieron la idea directamente de su existencia, o se derivaron de los testimonios de otros, y esos testimonios, a su vez, de otros testimonios... hasta que llegamos a aquellos que fueron testigos oculares y espectadores del acontecimiento.³

Me parece que este punto de vista debe llevarnos a la regresión infinita que he descrito más arriba. Porque el problema reside, por supuesto, en si "el testimonio unánime de los historiadores" ha de aceptarse, o si acaso ha de rechazarse como resultado de su dependencia de una fuente de información común, pero espuria. Recurrir a "letras presentes en nuestra memoria o en nuestros sentidos" no puede tener ningún peso en éste ni en otros problemas referentes a la historiografía.

Pero, entonces, ¿cuáles son las fuentes de nuestro conocimiento?

Pienso que la respuesta es como sigue: existe toda clase de fuentes de nuestro conocimiento; pero *ninguna tiene autoridad*.

Podemos afirmar que *The Times* puede ser una fuente de conocimiento, o que también puede serlo la *Encyclopaedia Britannica*. Podemos decir que ciertos artículos publicados en la *Physical Review* acerca de un problema de esta ciencia tienen más autoridad, y son más propiamente una fuente de conocimiento, que un artículo sobre el mismo problema publicado en *The Times* o en la *Encyclopaedia*. Pero sería un craso error considerar que la fuente del artículo publicado en la *Physical Review* debió de ser totalmente, o incluso parcialmente, observación pura. La fuente bien podría ser el descubrimiento de una incongruencia en otro artículo o, por ejemplo, el descubrimiento de que una hipótesis propuesta en otro artículo podría comprobarse con tal o cual experimento; todos estos descubrimientos "no observacionales" son "fuentes", en el sentido de que todos añaden algo a nuestro conocimiento.

No niego, por supuesto, que un experimento también puede añadir algo a nuestro conocimiento, y de manera importantísima. Pero de ninguna manera esencial constituye una fuente de conocimiento. Porque el experimento tiene que comprobarse siempre. Como en el ejemplo del periódico *The Times*, en general no preguntamos al testigo ocular de un experimento, sino que, si dudamos del resultado, podemos repetir el experimento o pedir a alguien que lo repita.

³ Véase D. Hume, *A Treatise of Human Nature [Tratado sobre la naturaleza humana]*, 1739, Primer Libro, 3ª parte, sección IV, edición de L. Selby-Bigge, p. 83. Véase también Hume, *Enquiry*, sección x; edición de Selby-Bigge, pp. 111 y ss.

El error fundamental que comete la teoría filosófica de las fuentes fundamentales de nuestro conocimiento es que no distingue con suficiente claridad entre preguntas de origen y preguntas de validez. Ciertamente es que, en el caso de la historiografía, estas dos clases de preguntas pueden a veces coincidir. La pregunta sobre la validez de una afirmación de índole histórica puede ser comprobable únicamente, o principalmente, a la luz del origen de ciertas fuentes de información. Pero, en general, las dos preguntas son de diferente clase; y en general no ponemos a prueba la validez de una aseveración o de una información rastreando sus fuentes o su origen, sino que la comprobamos mucho más directamente mediante el examen crítico de lo que se ha afirmado, de los hechos aseverados mismos.

Así pues, las preguntas del empirismo: "Cómo lo sabe usted?, ¿cuáles son las fuentes de su aserción?", están mal formuladas. Y no es que estén formuladas de manera inexacta o torpe; *están del todo erróneas en su concepción*: son preguntas que piden una respuesta autoritaria.

m

Podemos afirmar que los sistemas de epistemología tradicionales resultan de respuestas *sí* o de respuestas *no* a preguntas relativas a las fuentes de nuestro conocimiento. *Jamás ponen en tela de juicio estas preguntas, ni disputan sobre su legitimidad*: estas preguntas se consideran perfectamente naturales, y nadie parece ver en ellas nada perjudicial.

Esto es del mayor interés, porque estas preguntas surgen de un espíritu claramente autoritario. Podemos compararlas con aquella pregunta tradicional de la teoría política: "¿Quién debería gobernar?" que pide una respuesta autoritaria del género: "los mejores", o "los más sabios", o "el pueblo", o "la mayoría". (De paso, digamos que sugiere tan tontas opciones como "¿Quiénes deberían gobernar: los capitalistas, o los trabajadores?", análoga a "¿Cuál es la fuente esencial de nuestro conocimiento: el intelecto, o los sentidos?") Esta pregunta política está erróneamente formulada, y las respuestas que suscita son paradójicas [véase el escrito 25, más adelante]. Debería sustituirse con una pregunta del todo diferente, como "*¿Qué hacer, cómo organizar nuestras instituciones políticas para que los gobernantes malos o incompetentes (que debemos tratar de que no los haya, pero que tan fácilmente los hay todo el tiempo) no puedan hacer demasiado daño?*" Estoy convencido de que sólo cambiando nuestra pregunta en esta forma podemos aspirar a avanzar hacia una teoría razonable de las instituciones políticas.

La pregunta sobre las fuentes de nuestro conocimiento puede sustituirse de manera semejante a como lo hemos hecho con la anterior. Siempre se ha formulado en el espíritu de: "¿Cuáles son las mejores fuentes de nuestro conocimiento..., las más confiables, aquellas que no nos induzcan a error, y aquellas a las que podamos o debamos recurrir, en caso de duda, como a la corte suprema de apelaciones?" En vez de esto, propongo suponer que no existe tal fuente ideal de conocimiento —así como no existen los gobernantes ideales—, y que *todas* las "fuentes de conocimiento" a veces pueden inducirnos a error. Y, por tanto, propongo sustituir esta pregunta sobre las fuentes de nuestro conocimiento por esta otra, totalmente distinta: "*¿Cómo podemos esperar detectar y eliminar el error?*"

La pregunta acerca de las fuentes de nuestro conocimiento, como tantas otras autoritarias, es una pregunta *genética*. Nos pregunta sobre el origen de nuestro conocimiento, en la creencia de que el conocimiento puede legitimarse mediante su pedigrí. La nobleza del conocimiento racialmente puro, el conocimiento impoluto, el que deriva de la más alta autoridad, de ser posible de Dios: tales son las ideas metafísicas (a menudo inconscientes) que están detrás de la pregunta. De mi pregunta modificada: "¿Cómo podemos esperar detectar el error?", puede decirse que deriva del punto de vista de que tales fuentes de conocimiento tan puras, impolutas y certeras no existen, y que las preguntas referentes al origen o a la pureza de tales fuentes de conocimiento no deberían confundirse con las preguntas referentes a su validez, o a su veracidad. Podemos afirmar que este punto de vista es tan antiguo como Jenófanes. Éste sabía que nuestro conocimiento es conjetura, opinión —de *doxa*, más que de *epistémé*—, como se demuestra en sus versos [citados más arriba, en las páginas 30 y 31]. No obstante, la pregunta tradicional sobre las fuentes del conocimiento autorizadas se sigue repitiendo aún hoy, y la formulan muy a menudo algunos positivistas y otros filósofos que se consideran en rebeldía contra la autoridad.

La respuesta adecuada a mi pregunta de "¿Cómo esperamos detectar y eliminar el error?" es, en mi opinión, "*Criticando* las teorías o las conjeturas de los demás y —si podemos adiestrarnos a hacerlo así— *criticando* nuestras propias teorías o conjeturas". (Este último punto es muy deseable, pero no indispensable; porque si no somos capaces de criticar nuestras propias teorías, es posible que haya otros que lo hagan por nosotros.) Esta respuesta resume una posición que propongo llamar "racionalismo crítico". Constituye un punto de vista, una actitud y una tradición que debemos a los griegos. Es muy diferente del "racionalismo" o del "intelectualismo" de Descartes y su escuela, y muy diferente, incluso, de la epistemología de Kant. Sin embargo, en el campo de la

ética, del conocimiento moral, Kant se aproximó a este enfoque con su *principio de autonomía*. Este principio expresa su comprensión de que no debemos aceptar la orden de ninguna autoridad, por elevada que sea ésta, como base de la ética. Porque siempre que nos enfrentemos a la orden de una autoridad, a nosotros nos corresponde juzgar, críticamente, si es moral o inmoral obedecerla. La autoridad puede tener poder para vigilar el cumplimiento de sus órdenes, y acaso seamos impotentes para resistirnos a cumplirlas. Pero si tenemos el poder físico suficiente para elegir, la responsabilidad esencial recae en nosotros mismos. Radica en nuestra propia decisión crítica si debemos obedecer una orden; si hemos de someternos a la autoridad.

Kant aplicó osadamente esta idea al campo de la religión. Escribe: "[...] De cualquier modo, la Deidad debería darse a conocer a ti, e incluso [...] si se te revelara, eres tú [...] el único que debe juzgar si te está permitido creer en Ella, y si has de adorarla".⁴

En vista de esta valerosa declaración, parece extraño que en su filosofía de la ciencia Kant no haya adoptado la misma actitud de racionalismo crítico, de la búsqueda del error mediante la crítica. Estoy seguro de que fue sólo su aceptación de la autoridad de la cosmología de Newton —resultado de su casi increíble éxito al soportar las más severas pruebas— la que impidió a Kant aplicarle el racionalismo crítico. Si es correcta esta interpretación de Kant, entonces el racionalismo crítico (y también el empirismo crítico) que defendiendo sólo dio el toque final a la propia filosofía crítica de Kant. Y esto fue posible gracias a Einstein, quien nos enseñó que la teoría de Newton bien puede estar equivocada, a pesar de su éxito abrumador.

Así pues, mi respuesta a las preguntas "¿Cómo lo sabe usted?, ¿cuál es la fuente o la base de su afirmación?, ¿qué observaciones le han permitido llegar a ella?", sería: "No lo sé; mi aseveración era sólo una conjetura. No importa la fuente, no importan las fuentes, pues hay muchas fuentes de conocimiento y acaso no conozca yo ni la mitad de ellas; y sus orígenes o pedigrís, en todo caso, tienen poco que ver con la verdad. Pero, si le interesa a usted el problema que he intentado resolver con mi aseveración tentativa, me podrá ayudar criticándola tan severamente como pueda; y si puede diseñar alguna prueba experimental que considere refutaría mi aseveración, con todo gusto, y en la medida de mis posibilidades, le ayudaré a refutarla".

Esta respuesta se aplica, hablando en sentido estricto, sólo si la pregunta se refiere a alguna afirmación científica, distinta de una de carác-

⁴ Véase Immanuel Kant, *Religion within the Limits of Pure Reason* [La religión dentro de los límites de la razón pura], 2ª ed., 1794, capítulo IV, Segunda Parte, sección I, la primera nota al pie de página.

ter histórico. Si mi conjetura fue de índole histórica, las fuentes (en el sentido de no definitivas) por supuesto que tendrán que formar parte de la discusión crítica de su validez. Sin embargo, fundamentalmente, mi respuesta será la misma, como ya lo hemos visto.

IV

Creo que, llegados a este punto, es más que oportuno formular los resultados epistemológicos de esta discusión. Los expondré en la forma de nueve tesis:

1) No existen las fuentes esenciales de conocimiento. Cada fuente, cada sugerencia, es bienvenida; y cada fuente de información, cada sugerencia, está abierta al examen crítico. Excepto en historia, generalmente examinamos los hechos mismos, más que las fuentes de nuestra información.

2) La pregunta epistemológica adecuada no es acerca de las fuentes de información; en vez de ello, preguntamos si la aseveración que se ha hecho es verdadera..., es decir, si se apega a los hechos. (Que podemos funcionar, sin caer en antinomias, con la idea de la verdad objetiva en el sentido de su correspondencia con los hechos, lo ha demostrado la obra de Alfred Tarski.) Y tratamos de averiguar esta veracidad, tan bien como podamos, examinando o poniendo a prueba la aseveración misma; ya sea directamente, o examinando o poniendo a prueba sus consecuencias.

3) En relación con este examen, toda clase de argumentos pueden ser importantes. Un procedimiento típico es examinar si nuestras teorías concuerdan con nuestras observaciones. Pero también podemos examinar, por ejemplo, si nuestras fuentes históricas son mutua e internamente congruentes.

4) Cuantitativa y cualitativamente, con mucho, la fuente más importante de nuestro conocimiento —aparte del conocimiento innato— es la tradición. La mayoría de las cosas que sabemos las hemos aprendido mediante el ejemplo, porque nos las dijeron, por la lectura de libros, al aprender a hacer crítica, a aceptar y a hacer la crítica, y por aprender a respetar la verdad.

5) El hecho de que la mayoría de las fuentes de nuestro conocimiento sean tradicionales condena como fútil el antitradicionalismo. Pero este hecho no debe esgrimirse para apoyar una actitud tradicionalista; cada fragmento de nuestro conocimiento tradicional (y hasta nuestro conocimiento innato) está abierto al examen crítico, y puede descartarse. Sin embargo, sin la tradición, el conocimiento sería imposible.

6) El conocimiento no puede empezar desde la nada, desde hacer *tabula rasa* ni desde la observación. El avance del conocimiento consiste, principalmente, en la modificación del conocimiento previo. Aunque a veces podemos avanzar, por ejemplo en arqueología, mediante una observación fortuita, la importancia del descubrimiento dependerá generalmente de su poder para modificar nuestras teorías anteriores.

7) Las teorías epistemológicas pesimistas y optimistas son más o menos igualmente erróneas. El relato pesimista de la caverna de Platón es el que se apega a la verdad; no así su optimista relato de la *anamnesis* (aunque debemos aceptar que todos los hombres, como todos los demás animales, e incluso las plantas, poseen conocimiento innato). Mas aunque el mundo de las apariencias es ciertamente un mundo de meras sombras en las paredes de nuestra caverna, todos nos asomamos fuera de ella; y aunque, como dijo Demócrito, la verdad está oculta en las profundidades, podemos sondear las profundidades. No hay un criterio de la verdad a nuestra disposición, y este hecho apoya al pesimismo. Pero sí poseemos criterios que, *si tenemos suerte*, pueden permitirnos reconocer el error y la falsedad. Asimismo, la coherencia no puede establecer por sí misma la verdad, pero la incoherencia y la incongruencia sí establecen la falsedad. La claridad y la capacidad de distinguir no son criterios de la verdad, pero cosas como la oscuridad de expresión o la confusión *pueden* ser indicios de error. Y, cuando las reconocemos, nuestros propios errores nos proporcionan las tenues luces rojas que nos ayudan a encontrar a tientas el camino para salir de la oscuridad de nuestra caverna.

8) Ni la observación ni la razón constituyen una autoridad. La intuición intelectual y la imaginación son muy importantes, pero no son confiables; pueden mostrarnos cosas muy claramente, y sin embargo, desencaminarnos. Son indispensables como las fuentes principales de nuestras teorías; pero la mayoría de nuestras teorías son falsas de todos modos. La más importante función de la observación y del razonamiento, e incluso de la intuición y de la imaginación, es ayudarnos en el examen crítico de aquellas osadas conjeturas que constituyen los medios con los que sondeamos lo desconocido.

9) Cada solución de un problema da lugar a nuevos problemas no resueltos; cuanto más sea así, más profundo será el problema, y más osada su solución. Cuanto más aprendemos acerca del mundo, y cuanto más profundo sea nuestro aprendizaje, tanto más consciente, específico y articulado será nuestro conocimiento de lo que no sabemos, nuestro conocimiento de nuestra ignorancia. Porque esto, sin lugar a dudas, es la principal fuente de nuestra ignorancia: el hecho de que nuestro conocimiento sólo puede ser finito, mientras que nuestra ignorancia debe ser necesariamente infinita.

Podemos tener una vislumbre de la vastedad de nuestra ignorancia cuando contemplamos la vastedad del firmamento: aunque la sola magnitud del universo no es la causa más profunda de nuestra ignorancia, sí es una de sus causas. En un encantador pasaje de su obra, F. P. Ramsey escribió:

En lo que parezco estar en desacuerdo con mis amigos es en atribuir poca importancia al tamaño físico. No siento la menor humildad ante la vastedad de los cielos. Por más que sean enormes las estrellas, no pueden pensar ni amar; y éstas son cualidades que me impresionan más que el tamaño. No me parece un gran mérito pesar cerca de cien kilos.⁵

Sospecho que los amigos de Ramsey habrían estado de acuerdo con él respecto a la importancia del tamaño de las cosas; y también sospecho que, si se sentían humildes ante la vastedad del universo, era así porque veían en ello un símbolo de su ignorancia.

Estoy convencido de que valdría la pena tratar de aprender algo acerca del mundo, aun si en este intento sólo nos enteráramos de que no sabemos gran cosa acerca de él. Este estado de ilustrada ignorancia podría ayudarnos en muchos de nuestros conflictos. Sería conveniente que todos entiendiéramos que, aunque diferimos en gran medida en cuanto a los insignificantes chispazos de conocimiento que tenemos, en lo que respecta a nuestra infinita ignorancia, todos somos iguales.

Hay un último asunto que me gustaría tratar aquí.

Si nos tomamos el trabajo de buscarla, a menudo podemos encontrar una idea verídica, merecedora de que se le conserve, en una teoría filosófica que deba rechazarse como falsa. ¿Podemos encontrar una idea como ésta en una de las teorías de las fuentes originarias de nuestro conocimiento?

Creo que sí, y sugiero que es precisamente una de las dos ideas principales que subyacen en la doctrina de que la fuente de todo nuestro conocimiento reside en lo sobrenatural. La primera de estas ideas es falsa, en mi opinión, en tanto que la segunda es verídica.

La primera idea, la que considero falsa, es que debemos *justificar* nuestro conocimiento, o nuestras teorías, con razones *positivas*; esto es, con razones capaces de afirmarlas, o por lo menos de hacerlas muy probables;

⁵ Véase F. P. Ramsey, *The Foundations of Mathematics* [Los fundamentos de las matemáticas], 1931, p. 291.

en cualquier caso, por mejores razones que éstas, hasta ahora han soportado bien la crítica. Esta idea implica, sugiero, que debemos recurrir a alguna originaria o autoritaria fuente de verdadero conocimiento; lo cual deja abierto el carácter de esa autoridad... ya sea humana, como la observación o la razón, o sobrehumana (y por tanto sobrenatural).

La segunda idea —cuya vital importancia ha recalcado Russell— es que ninguna autoridad humana puede establecer la verdad por decreto: que debemos todos someternos a la verdad; que *la verdad está por encima de toda autoridad humana*.

Consideradas juntas, estas dos ideas casi inmediatamente suscitan la conclusión de que las fuentes de las que deriva nuestro conocimiento deben de ser sobrehumanas; conclusión que tiende a alentar el fari-seísmo y el empleo de la fuerza contra los que se niegan a ver la verdad divina.

Algunos que acertadamente rechazan esta conclusión no rechazan también, por desgracia, la primera idea: la creencia en la existencia de las fuentes fundamentales del conocimiento. En vez de ello, rechazan la segunda idea: la tesis de que la verdad está por encima de toda autoridad humana. Por tanto, así ponen en peligro la idea de la objetividad del conocimiento, y la de que debe haber normas comunes de crítica o racionalidad.

Sugiero que lo que debemos hacer es renunciar a la idea de las fuentes originarias del conocimiento, y aceptar que todo conocimiento humano es humano: que está mezclado con nuestros errores, con nuestros prejuicios, nuestros sueños y nuestras esperanzas: que lo único que podemos hacer es buscar a tientas la verdad, aunque esté más allá de nuestro alcance. Podemos admitir que nuestro esfuerzo por alcanzarla a veces resulta inspirado, pero debemos estar en guardia contra la creencia, por más hondamente que la sintamos, de que nuestra inspiración está investida de alguna clase de autoridad, divina o de otra índole. Si aceptamos que no existe ninguna autoridad más allá del alcance de la crítica que se encuentre dentro del campo de nuestro conocimiento, por muy lejos que haya penetrado en lo desconocido, entonces sí podremos retener, sin peligro, la idea de que la verdad está más allá de la autoridad humana. Y debemos retener esta idea. Porque sin esta idea no puede haber normas objetivas de búsqueda de la verdad; sin ella, no habrá crítica para nuestras investigaciones; no habrá sin ella ese buscar a tientas lo desconocido; sin ella, no habrá búsqueda del conocimiento.

4. CONOCIMIENTO: SUBJETIVO CONTRA OBJETIVO (1967)

I. TRES TESIS SOBRE LA EPISTEMOLOGÍA DEL MUNDO NÚMERO 3

PROBABLEMENTE haya yo escandalizado a quienes hayan oído hablar de mi actitud adversa hacia Platón y hacia Hegel cuando titulé mi conferencia "Una teoría del mundo platónico", o "Una teoría del espíritu objetivo".

El tema principal de esta conferencia será lo que a menudo llamo, a falta de un nombre mejor, el "mundo número 3". Para explicar esta expresión, señalaré que, sin tomar demasiado en serio las palabras "mundo" o "universo", podemos distinguir entre los siguientes tres mundos o universos: primero, el mundo de los objetos físicos o de los estados físicos; segundo, el mundo de los estados de conciencia, o estados mentales o, quizá, de las disposiciones conductuales para actuar; y tercero, el mundo de *los contenidos objetivos del pensamiento*, especialmente el de los pensamientos científico y poético y el de las obras de arte.

Así, lo que llamo el "mundo número 3" tiene, por cierto, mucho en común con la teoría platónica de las Formas o las Ideas y, por tanto, también con el Espíritu Objetivo de Hegel, aunque mi teoría difiere radicalmente, en algunos aspectos decisivos, de la de Platón y de la de Hegel. Tiene más en común con la teoría de Bolzano, de un universo de proposiciones en sí mismas y de verdades en sí mismas, aunque también difiere de la de Bolzano. Mi mundo número 3 se parece mucho más al universo de Frege, de los contenidos objetivos del pensamiento.

No forma parte de mi punto de vista ni de mi argumento que no podamos enumerar nuestros mundos de maneras diferentes, o que no los enumeremos en absoluto. Podríamos distinguir, especialmente, entre más de tres mundos. Mi expresión "mundo número 3" es simplemente un término que empleo por conveniencia.

Al preconizar un "tercer mundo" objetivo, espero provocar reacciones en aquellos a los que llamo "filósofos de creencias", aquellos que, como Descartes, Locke, Berkeley, Hume, Kant o Russell, están interesados en nuestras creencias subjetivas, y atacar sus bases de origen. Contra estos filósofos de creencias deseo enfatizar que nuestro problema consiste en mejores y más audaces teorías; y que en esto cuenta la *preferencia crítica*, pero no las *creencias*.

También deseo confesar, no obstante, desde el principio, que soy realista: sugiero, en mi calidad de algo así como un realista ingenuo, que existen un mundo físico (mundo número 1), un mundo de estados de conciencia (mundo número 2), y que entre estos dos existe una interacción. Y estoy convencido de que existe un tercer mundo, en un sentido que a continuación explicaré más detalladamente.

Entre los habitantes de mi "mundo número 3" están, más especialmente, los *sistemas teóricos*; pero otros habitantes igualmente importantes de él son los *problemas* y las *situaciones problemáticas*. Y argüiré que los más importantes habitantes de este mundo son los *argumentos de crítica*, y lo que puede llamarse —por analogía con un estado físico o con un estado de conciencia— *el estado de una discusión* o *estado de un argumento de crítica*; y, por supuesto, el contenido de publicaciones periódicas, libros y bibliotecas.

La mayoría de los que se oponen a la tesis de que hay un mundo objetivo número 3 aceptarán, por supuesto, que existen problemas, conjeturas, teorías, argumentos, revistas y periódicos y libros. Pero objetan generalmente que todas estas entidades son, esencialmente, *expresiones simbólicas* o *lingüísticas* de estados mentales subjetivos, o quizá de disposiciones conductuales para actuar; objetan, además, que estas entidades son medios de *comunicación*; es decir, medios simbólicos o lingüísticos para evocar en los otros hombres estados mentales semejantes o disposiciones conductuales para actuar.

Contra esto, he alegado a menudo que no es posible relegar todas estas entidades y su contenido al mundo número 2.

Permítaseme repetir uno de mis argumentos clásicos¹ en favor de *la existencia* (más o menos) *independiente del mundo número 3*.

Considero dos experimentos de pensamiento:

Experimento 1: Todas nuestras máquinas y herramientas son destruidas, y todo nuestro aprendizaje objetivo, incluso el conocimiento subjetivo de las máquinas y de las herramientas, y de su uso, han desaparecido. Pero sobreviven *las bibliotecas* y *nuestra capacidad para aprender de ellas*. Está claro que, en estas circunstancias, tras mucho sufrimiento, nuestro mundo podrá seguir adelante.

Experimento 2: Como en el anterior experimento hipotético, se han destruido todas las máquinas y herramientas, y también nuestro aprendizaje subjetivo, incluido nuestro conocimiento subjetivo de las máquinas y de las herramientas, y de la forma de usarlas. Pero esta vez *también han quedado destruidas todas las bibliotecas*, de modo que se ha vuelto inútil nuestra capacidad para aprender de los libros.

¹ Se ha adaptado el argumento de *La sociedad abierta y sus enemigos*, capítulo 15, sección m.

Si piensa usted en estos dos experimentos hipotéticos, la realidad, la importancia y el grado de autonomía del mundo número 3 (así como sus efectos en los mundos número 1 y número 2) quizá se le aclaren un tanto. Porque en el segundo caso no habrá resurgimiento de nuestra civilización durante muchos milenios.

Es mi propósito defender en esta conferencia tres tesis principales, las cuales se refieren a la epistemología. Considero que la epistemología es la teoría del *conocimiento científico*.

Mi primera tesis es ésta: La epistemología tradicional ha estudiado el conocimiento o el pensamiento en un sentido subjetivo: en el sentido del empleo ordinario de las frases "Yo sé, o conozco" o "Estoy pensando". Esto, lo asevero, ha hecho caer a los estudiantes de epistemología en incongruencias o en conceptos fuera de lugar: al pretender estudiar el conocimiento científico, en realidad estudiaban algo que no tiene nada que ver con el conocimiento científico. Porque el *conocimiento científico* simplemente no es conocimiento en el sentido del empleo ordinario que damos a la frase "Yo sé". Mientras que el conocimiento en el sentido de "Yo sé" pertenece a lo que yo llamo el "mundo número 2", el mundo de los *sujetos*, el conocimiento científico pertenece al mundo número 3, al mundo de las *teorías objetivas*, de los problemas objetivos, de los argumentos objetivos.

Así, mi primera tesis es que la epistemología tradicional, la de Locke, Berkeley, Hume, y hasta la de Russell, está fuera de lugar, en un sentido bastante estricto de la palabra epistemología. Un corolario de esta tesis es que una gran parte de la epistemología contemporánea también está fuera de lugar.

Mi primera tesis implica la existencia de dos diferentes significados de la palabra conocimiento o de la palabra pensamiento: 1) *conocimiento o pensamiento en el sentido subjetivo*, que consiste en un estado mental o de conciencia, o en la disposición para comportarse o reaccionar, y 2) *conocimiento o pensamiento en el sentido objetivo*, que consta de problemas, teorías y argumentos como tales. El conocimiento, en este sentido objetivo, es del todo independiente de la pretensión de alguien de saber; también es independiente de la creencia de alguien, o de su disposición a asentir, a afirmar algo o a actuar. El conocimiento, en el sentido objetivo, es *conocimiento sin cognosciente*: es *conocimiento sin sujeto que conoce*.

Frege escribió acerca del pensamiento en el sentido objetivo: "Entiendo por un *pensamiento* no el acto subjetivo de pensar, sino su contenido objetivo..."²

² Véase la p. 32 de G. Frege, "Über Sinn und Bedeutung", en *Zeitschrift für Philosophie und Philosophische Kritik* 100, 1892, pp. 25-50; las itálicas son mías. (Este ensayo de Frege

Los dos sentidos del concepto *pensamiento* y sus interesantes interrelaciones pueden ilustrarse con la siguiente cita, muy convincente, de Heyting, quien dice acerca del acto de Brouwer de inventar su teoría del *continuum*.³ "Si se hubieran inventado antes las funciones recursivas [que pueden arbitrar recursos], él [Brouwer] quizá no habría formulado la noción de una secuencia de opción, lo cual, en mi opinión, habría sido desafortunado."

Esta cita se refiere, por una parte, a algunos *procesos de pensamiento subjetivo* de Brouwer, y dice que podrían no haber ocurrido (lo cual habría sido desafortunado) si hubiera sido diferente la *situación problemática objetiva*. Así, Heyting menciona ciertas posibles *influencias* en los procesos de pensamiento subjetivo de Brouwer, y también expresa su opinión sobre el valor de estos procesos de pensamiento subjetivo. Ahora bien, es interesante comprobar que las influencias, *qua* influencias [como influencias en sí mismas, como tales], deben ser subjetivas: sólo el conocimiento subjetivo de Brouwer de las funciones recursivas habría podido tener ese efecto negativo de impedirle inventar las secuencias de opción libre.

Por otra parte, la cita de Heyting señala cierta relación objetiva entre los *contenidos objetivos* de dos pensamientos o teorías: Heyting no se refiere a las condiciones subjetivas o a la electroquímica de los procesos cerebrales de Brouwer, sino a una *situación problemática objetiva en matemáticas* y a sus posibles influencias en los actos de pensamiento subjetivo de Brouwer, que fueron soslayadas al resolver estos problemas objetivos. Yo describiría esto diciendo que la observación de Heyting se refiere a la *lógica situacional* objetiva, o del "mundo número 3" de la invención de Brouwer, y que la observación de Heyting implica que la situación del mundo número 3 puede afectar al mundo número 2. De manera similar, la sugerencia de Heyting de que habría sido desafortunado que Brouwer no hubiera inventado las secuencias de elección es otro modo de decir que el *contenido objetivo* del pensamiento de Brouwer era valioso e interesante: valioso e interesante, es decir, en la manera como cambió la situación del problema objetivo en el mundo número 3.

Para expresar este asunto en forma llana, si digo: "El pensamiento de Brouwer fue influido por Kant", o que "Brouwer rechazó la teoría del espacio de Kant", entonces estaré hablando por lo menos parcialmente

se ha traducido al inglés y reproducido en varias publicaciones; por ejemplo, por H. Feigl, en las pp. 85-102 de H. Feigl y W. Sellars (comps.), *Readings in Philosophical Analysis*, 1949.)

³ Véase la p. 195 de A. Heyting, "After Thirty Years" ["Treinta años después"], pp. 194-197 de E. Nagel, P. Suppes y A. Tarski (comps.), *Logic, Methodology and Philosophy of Science*, 1962.

de actos de pensamiento, en el sentido subjetivo: la palabra *influencia* indica un contexto de procesos de pensamiento o de actos de pensar. Sin embargo, si digo: "El pensamiento de Brouwer difiere mucho del de Kant", entonces está claro que hablo principalmente de contenidos. Y, por último, si digo: "Los pensamientos de Brouwer son incompatibles con los de Russell", entonces estaré utilizando un *término lógico: incompatible*, y con esto expreso sin ambigüedad alguna que estoy utilizando el término "pensamiento" sólo en el sentido objetivo que le atribuye Frege, y que estoy hablando sólo acerca del contenido objetivo, o del contenido lógico, de teorías.

Así como el lenguaje común no tiene, por desgracia, dos términos distintos para expresar "pensamiento" en el sentido del mundo número 2 y en el sentido del mundo número 3, tampoco tiene términos distintos para los correspondientes dos sentidos de "Yo sé" y de "conocimiento".

Para demostrar que ambos sentidos existen, mencionaré primero tres ejemplos subjetivos pertenecientes al mundo número 2:

J) "Yo sé que usted está tratando de provocarme, pero no dejaré que me provoque."

2) "Yo sé que el último teorema de Fermat no se ha comprobado, pero creo que algún día se comprobará."

3) Según la entrada "Conocimiento" del *Oxford English Dictionary*, *conocimiento* es "un estado de estar enterado o informado"

A continuación mencionaré tres ejemplos objetivos del mundo número 3:

1) Según la entrada "Conocimiento" del *Oxford English Dictionary*, el *conocimiento* es también "una rama del aprendizaje; una ciencia; un arte".

2) "Tomando en cuenta el estado actual del conocimiento *metamatemático*, parece que el último teorema de Fermat no se puede elucidar, por ahora."

3) "Certifico que esta tesis es una original e importante *contribución al conocimiento*."

Estos muy triviales ejemplos sólo tienen la función de ayudar a esclarecer lo que quiero decir cuando hablo de "conocimiento en el sentido objetivo". Mi cita de la entrada del *Oxford English Dictionary* no debe interpretarse ni como una concesión al análisis del lenguaje ni como un intento de apaciguar a sus partidarios. Tampoco he citado ese artículo con la intención de probar que "el empleo ordinario" del lenguaje abar-

ca la definición de "conocimiento" en el sentido objetivo de mi mundo número 3. En realidad, me sorprendió encontrar en el *Oxford English Dictionary* empleos objetivos de la palabra "conocimiento". (Me asombró más aún encontrar allí algunos, por lo menos *parcialmente* objetivos, empleos del verbo "conocer": "distinguir..., estar enterado o estar al tanto de (una cosa, un lugar, una persona)...; entender".) En todo caso, mis ejemplos no pretenden ser argumentos. Sólo intentan ser ilustraciones de mi tesis.

Mi *primera tesis*, que hasta este momento no he apoyado con argumentos, sino sólo con ilustraciones, era que la epistemología tradicional, con su concentración en el mundo número 2, o con su concepto del conocimiento situado en el sentido subjetivo, es ajena al estudio del conocimiento científico.

Mi *segunda tesis* consiste en afirmar que lo propio de la epistemología es el estudio de los problemas científicos y de las situaciones problemáticas, el estudio de las conjeturas científicas (que para mí es otra expresión para referirme a las hipótesis o teorías científicas), de las discusiones de índole científica, de los argumentos críticos y del papel que desempeña la evidencia en los argumentos; y, por tanto, de las publicaciones periódicas científicas y de los libros, y de los experimentos y su evaluación en la argumentación científica; o, en pocas palabras, que el estudio de un mundo número 3 *en gran medida autónomo* de conocimiento objetivo tiene una importancia decisiva para la epistemología.

Un estudio epistemológico, como el que he descrito en mi segunda tesis, demuestra que los científicos, con mucha frecuencia, no pretenden que sus conjeturas sean verdaderas, o que las "sepan" en el sentido subjetivo de "saber o conocer", o que crean en ellas. Aunque en general no afirman conocer, al llevar a cabo sus programas de investigación actúan basados en conjeturas acerca de lo que es o no fructífero, y en qué línea de investigación promete mejores resultados en el mundo número 3, el mundo del conocimiento objetivo. En otras palabras, los científicos actúan basándose en una conjetura o, si ustedes quieren, en una *creencia subjetiva* (pues así podemos llamar a la base subjetiva de una acción) respecto a lo que es prometedor de inminente *crecimiento en el mundo número 3, el mundo del conocimiento objetivo*.

Esto, sugiero, nos proporciona un argumento en favor, tanto de mi *primera tesis* (la de la irrelevancia de una epistemología subjetivista), como de mi *segunda tesis* (la de la relevancia de una epistemología objetivista).

Pero tengo una *tercera tesis*. Hela aquí: Una epistemología objetivista que estudie el mundo número 3 puede ayudarnos a arrojar una gran luz

sobre el mundo número 2, el mundo de la conciencia subjetiva, especialmente sobre los procesos de pensamiento subjetivo de los científicos; pero *lo contrario no es verdad*.

Así quedan formuladas mis tres tesis principales.

Además de mis tres tesis principales, presento a continuación otras tres de apoyo a ellas.

La primera es que el mundo número 3 es un producto natural del animal humano, comparable a la tela de la araña.

La segunda tesis auxiliar (y es una tesis casi decisiva, en mi opinión) consiste en que el mundo número 3 es en gran medida autónomo, aunque actuamos constantemente sobre él y aunque él actúa constantemente sobre nosotros; es autónomo, pese a que es producto nuestro y aunque tiene un fuerte efecto de retroalimentación sobre nosotros; es decir, sobre nosotros, en tanto que habitantes del mundo número 2 e incluso en tanto que habitantes del mundo número 1.

La tercera tesis auxiliar consiste en que mediante esta interacción entre nosotros y el mundo número 3 es como crece el conocimiento objetivo, y en que existe una muy estrecha analogía entre el crecimiento cognoscitivo y el crecimiento biológico: en otras palabras, entre el crecimiento de nuestro saber y la evolución de plantas y animales.

II. UNA APROXIMACIÓN BIOLÓGICA AL MUNDO NÚMERO 3

En esta sección trataré de defender la existencia de un mundo número 3 autónomo mediante un argumento de índole biológica o evolutiva.

El biólogo acaso se interese por estudiar la conducta animal; pero también puede estar interesado en las *estructuras no vivientes* que producen los animales, tales como las telarañas, los nidos que construyen las avispas o los que construyen las hormigas, las madrigueras subterráneas de los tejones, las represas que construyen los castores o las sendas que hacen los animales del bosque.

Distinguiré entre dos categorías principales de problemas que surgen del estudio de estas estructuras. La primera consta de problemas relacionados con *los métodos que utilizan los animales* o con *las maneras de comportarse de los animales* cuando construyen estas estructuras. Por tanto, esta primera categoría abarca *los problemas relacionados con los actos de producción*; las disposiciones conductuales del animal, y las relaciones entre el animal y el producto. La segunda categoría de problemas se ocupa de las *estructuras mismas*. Se ocupa de la química de los materiales utilizados en la estructura; de sus propiedades geométricas y físi-

cas; de sus cambios evolutivos, dependientes de especiales condiciones ambientales, y de su dependencia de estas condiciones ambientales o de su adaptación a ellas. Muy importante también es la *relación de retroalimentación* surgida de las propiedades de la estructura con la conducta de los animales. Al tratar con los problemas de esta segunda categoría —es decir, con las estructuras mismas—, también tendremos que ocuparnos de las estructuras desde el punto de vista de sus funciones biológicas. Así, algunos problemas de la primera categoría tendrán que surgir cuando discutamos problemas de la segunda categoría; por ejemplo: "¿Cómo fue construido este nido?" y "¿Qué aspectos de su estructura son típicos (y por ello probablemente tradicionales o heredados) y qué aspectos son variantes adaptadas a condiciones especiales?"

Como se ve claramente en mi último ejemplo de problema, los problemas de la primera categoría —esto es, los relacionados con la producción de la estructura— a veces nos los sugerirán los problemas de la segunda categoría. Y es lógico que así sea, puesto que ambas categorías de problemas dependen *del hecho de que tales estructuras objetivas existen*, hecho que, en sí mismo, pertenece a la segunda categoría. Así pues, podemos decir que las *estructuras mismas* crean ambas categorías de problemas. Podemos decir que la segunda categoría de problemas —los relacionados con las estructuras mismas— es la más fundamental: lo único que nos hace presuponer de la primera categoría es el hecho mismo de que las estructuras son *producidas de algún modo* por algunos animales.

Ahora bien, estas sencillas consideraciones también pueden aplicarse, por supuesto, a productos de la actividad *humana*, tales como casas o herramientas, y también a las obras de arte. Y lo que es más importante para nosotros: se aplican a lo que llamamos "lenguaje" y a lo que llamamos "ciencia".⁴

La conexión entre estas consideraciones biológicas con el tema de mi presente conferencia puede aclararse reformulando mis tres tesis principales. Mi primera tesis puede expresarse diciendo que, en la situación actual de los problemas de la filosofía, pocas cosas son tan importantes como el darse cuenta de la distinción entre las dos categorías de problemas: problemas de la producción, por un lado, y, por el otro, problemas relacionados con las estructuras producidas mismas. Mi segunda tesis es que deberíamos darnos cuenta de que la segunda categoría de problemas, los relacionados con los productos en sí mismos, es en casi todos los aspectos más importante que los de la primera categoría, los

⁴ Sobre estos "artefactos", véase la p. 111 de F. A. von Hayek, *Studies in Philosophy, Politics and Economics*, 1967.

problemas de la producción. Mi tercera tesis es que los problemas de la segunda categoría son básicos para entender los problemas de la producción: al contrario de lo que podría parecer a las primeras impresiones, podemos aprender más sobre la conducta de producción estudiando los productos mismos que lo que podemos aprender acerca de los productos estudiando la conducta de producción. Esta tercera tesis puede describirse como anticonductualista y antipsicologista.

En su aplicación a lo que puede llamarse "conocimiento", mis tres tesis pueden formularse como sigue:

1) Deberíamos tener presente en todo momento la distinción entre problemas relacionados con nuestra contribución personal a la producción de conocimiento científico, por una parte, y, por la otra, problemas relacionados con la estructura de los diversos productos, tales como teorías científicas o argumentos científicos.

2) Deberíamos advertir que el estudio de los productos es mucho más importante que el estudio de la producción, incluso para entender cabalmente la producción y sus métodos.

3) Podemos aprender más acerca de la heurística y de la metodología, e incluso acerca de la psicología de la investigación, estudiando teorías y los argumentos que nos presentan en favor o en contra de estas teorías, que mediante algún enfoque conductualista, psicológico o sociológico. En general, podemos aprender muchísimo acerca de la conducta y de la psicología estudiando directamente los productos.

En lo que sigue, llamaré al acercamiento desde los productos —tanto las teorías como los argumentos— el enfoque "objetivo", o el enfoque desde el punto de vista del "mundo número 3". Y llamaré a los enfoques conductualista, psicológico y sociológico del conocimiento científico el enfoque "subjetivo", o enfoque desde el punto de vista del "mundo número 2".

El atractivo que ejerce el enfoque subjetivo se debe en gran medida a que es *causal*. Porque acepto que las estructuras objetivas a las que atribuyo prioridad son causadas por la conducta humana. Y como es causal, el enfoque subjetivo puede parecer más científico que el enfoque objetivo, el cual, por así decirlo, empieza a operar a partir de efectos, en vez de a partir de causas.

Aunque reconozco que las estructuras objetivas son productos de la conducta, sostengo que el argumento es erróneo. En todas las ciencias, el enfoque ordinario es partir de los efectos hacia las causas. El efecto suscita el problema —el problema que hay que explicar, el *explicandum*—, y el científico intenta resolverlo elaborando una hipótesis explicativa.

Mis tres tesis principales, con el hincapié en los productos objetivos, no son, por tanto, ni ideológicas ni se apartan de lo científico.

III. LA OBJETIVIDAD Y LA AUTONOMÍA DEL MUNDO NÚMERO 3

Una de las razones principales para adoptar el enfoque subjetivo del conocimiento es considerar que un libro no es nada sin el lector: sólo si se entiende el libro se convierte realmente en un libro; si no se entiende, es sólo papel con manchas negras.

Este punto de vista es erróneo en muchos aspectos. Un nido de avispas sigue siendo un nido de avispas aun después de haber quedado vacío; aunque las avispas no vuelvan a usarlo como nido. Un nido de ave sigue siéndolo aunque ningún pájaro haya vivido en él. De modo semejante, un libro sigue siendo un libro —cierto tipo de producto— aunque nunca lo lea nadie (como puede suceder fácilmente en nuestros días).

Además, puede ocurrir que un libro, o incluso toda una biblioteca, no necesite que alguien lo haya escrito, en forma creativa: una serie de libros de logaritmos, por ejemplo, puede producirse e imprimirse mediante computadora. Acaso sea el mejor libro de logaritmos; acaso contenga logaritmos, digamos, hasta de cincuenta decimales. Puede enviarse a nuestras librerías, pero quizá se considere muy incómodo para consultarse; en todo caso, pueden pasar años antes de que alguien lo utilice, y muchas de sus cifras (que representan teoremas matemáticos) acaso nadie las vea mientras haya hombres sobre la faz de la tierra. Sin embargo, cada una de estas cifras contiene lo que yo llamo "conocimiento objetivo", y la pregunta de si tengo o no derecho a llamarlo así no tiene ningún interés.

El ejemplo de estos libros de logaritmos puede parecer muy artificioso, pero no lo es. Podría yo decir que casi todo libro es así: contiene conocimiento objetivo, falso o verdadero, útil o inútil; y el que nadie lo lea nunca y nadie entienda en realidad su contenido resulta casi accidental. El hombre que lee un libro y lo entiende perfectamente es una rara creatura. Pero aunque hubiera muchos de estos hombres, siempre habría en ellos mucha falta de comprensión y no pocas malas interpretaciones; y no es el evitar realmente y en forma un tanto accidental tales malas interpretaciones lo que convierte a ese conjunto de manchas negras sobre papel blanco en un libro, o sea, una instancia del conocimiento, en el sentido objetivo. Se trata, más bien, de algo más abstracto. Es su posibilidad o potencialidad de que se le entienda y se le interprete bien, su carácter disposicional de ser bien entendido y bien interpretado, o a que no se le entienda y se le interprete mal, lo que convierte una cosa en un libro. Y esta potencialidad o disposición puede existir sin que siquiera se haya actualizado o sin que nadie se haya dado cuenta de ello.

Para ver esto más claramente, podemos imaginar que, después de que la raza humana haya perecido, algunos civilizados sucesores nuestros encontrarán algunos libros o bibliotecas (no importa que estos sucesores nuestros sean animales terrícolas que se hayan vuelto civilizados, o algunos visitantes procedentes del espacio extraterrestre). Estos libros pueden descifrarse. Podría tratarse de esas tablas de logaritmos que jamás se hubieran leído, para reforzar nuestro argumento. Esto esclarece bien que ni su composición por animales pensantes, ni el hecho de que en realidad no se hayan leído o entendido, son esenciales para que una cosa, un objeto, sea un libro, y que basta con que pudiera descifrarse.

Así, reconozco que, para pertenecer al mundo número 3, el mundo del conocimiento objetivo, un libro debería ser capaz —en principio, o virtualmente— de ser captado (o descifrado, o entendido, o "sabido", o conocido) por alguien. Pero no hago ninguna otra concesión al respecto.

Podemos afirmar, por lo antedicho, que existe una clase de mundo número 3 de libros como tales, platónicos (o bolzanescos), de teorías como tales, de problemas como tales, de situaciones problemáticas como tales, de argumentos como tales, etc. Y asevero que, aunque este mundo número 3 es un producto humano, hay muchas teorías como tales, muchos argumentos como tales y situaciones problemáticas como tales que nunca se han producido o entendido y que acaso los hombres nunca los produzcan o entiendan.

La tesis de la existencia de tal mundo de situaciones problemáticas les parecerá a muchos extremadamente metafísica y dudosa, pero puede defenderse al señalar su análoga tesis biológica. Por ejemplo, tiene su analogía incuestionable en el ámbito de los nidos de las aves. Hace algunos años me regalaron algo para mi jardín: una caja para anidación de pájaros. Era un producto humano, por supuesto, y no un producto de ave, así como nuestro libro de tablas de logaritmos era un producto de computadora más que un producto humano. Pero, en el contexto del mundo de las aves, era parte de una situación problemática objetiva, y una oportunidad objetiva. Los pájaros, durante algunos años, no parecían haber notado siquiera la existencia de esa caja para anidar. Pero, al cabo de algunos años, unos paros azules la inspeccionaron cuidadosamente, e incluso empezaron a construir un nido en ella, y al poco tiempo dejaron de hacerlo. Era obvio que había allí una oportunidad detectable, aunque, al parecer, no muy valiosa en especial. El caso es que había allí una situación problemática. Y el problema pueden resolverlo otros pájaros, en otro año. Si no es así, otra caja para anidar acaso sea más apropiada. Por otra parte, la caja más apropiada puede quitarse de allí antes de que se utilice. La cuestión de lo apropiado de la caja es,

claramente, una cuestión objetiva; y el que la caja sea utilizada alguna vez resulta parcialmente accidental. Y esto es lo que ocurre en todos los nichos ecológicos. Son potencialidades, y pueden estudiarse como tales de manera objetiva, hasta cierto punto independientes de la pregunta de si estas potencialidades las actualizará alguna vez un organismo viviente. El bacteriólogo sabe preparar un nicho ecológico de este género, para cultivar ciertas bacterias o ciertos mohos. Acaso este nicho o caldo de cultivo sea el apropiado para su propósito. Sin embargo, el que alguna vez lo utilicen o habiten ciertos organismos, eso ya es otra cosa muy distinta.

Una gran parte del mundo objetivo número 3 de teorías y libros y argumentos potenciales surge como un subproducto no intencional de los libros y argumentos que en realidad se producen. Podemos afirmar, también, que constituye un subproducto del lenguaje humano. El lenguaje mismo, como un nido de ave, es un subproducto no intencional de acciones que estaban dirigidas hacia otros propósitos.

¿Cómo surge en la selva un sendero hecho por animales? Algunos animales acaso irrumpen en la maleza para tener acceso a un abrevadero. Otros animales ven que es fácil utilizar esa senda abierta. Así, puede ampliarse y mejorarse mediante el uso. No es algo planificado: constituye una consecuencia no intencional de la necesidad de disponer de desplazamiento fácil y rápido. Así es como originalmente se hace una senda —quizá también sea así como la hacen incluso los hombres—, y así puede ser el surgimiento del lenguaje y de otras instituciones humanas útiles, y acaso deban su existencia y su desarrollo a la utilidad de estas instituciones. No estaban planificadas ni se tenía la intención de crearlas, y acaso no se necesitaran antes de que llegaran a existir; pero pueden crear una nueva necesidad, o un nuevo conjunto de metas o propósitos: la estructura orientada hacia la consecución de metas de los animales o de los hombres no es algo que esté "dado", sino que se desarrolla, con la ayuda de algún mecanismo de retroalimentación, a partir de metas previas, y a partir de resultados que se buscaba o no se buscaba lograr.⁵

De esta manera, puede surgir todo un nuevo universo de posibilidades o de potencialidades: un mundo que, en gran medida, es *autónomo*.

Un ejemplo muy claro de esto es un jardín. Aunque se haya planificado con sumo cuidado, en general crecerá en parte de maneras inesp-

⁵ Véase Hayek, *op. cit.*, capítulo 6, especialmente las pp. 96 y 100, nota 12; Rene Descartes, *Discourse on Method [Discurso del método y meditaciones metafísicas]*, edición inglesa de E. S. Haldane y G. R. T. Ross, p. 89 [escrito 24, sección II, más adelante]; y *Objective Knowledge [Conocimiento objetivo]*, pp. 253-255.

radas. Pero, aunque crezca tal como se planificó, algunas inesperadas interrelaciones entre los objetos planificados pueden hacer surgir todo un universo de posibilidades, de posibles nuevas metas y de nuevos *problemas*.

El mundo del lenguaje, el de las conjeturas, de las teorías y de los argumentos —en pocas palabras, el universo del conocimiento objetivo—, es uno de los más importantes de estos universos creados por el hombre, mas al mismo tiempo, en gran parte, *autónomos*.

La idea de *autonomía* es central en mi teoría del mundo número 3: si bien el mundo número 3 es un producto humano, creación humana, crea a su vez, como lo hacen otros productos animales, su propio *dominio de autonomía*.

Hay incontables ejemplos de esto. Quizá los más impresionantes y, en todo caso, los que deberíamos recordar siempre como nuestros ejemplos estándares, se encuentren en la teoría de los números naturales.

Que Kronecker *se apacigüe*: convengo con Brouwer en que la secuencia de los números naturales constituye una construcción humana. Pero aunque hemos creado esta secuencia, ésta crea, a su vez, sus propios problemas autónomos. La distinción entre números pares y números impares no ha sido creada por nosotros: es una consecuencia no intencional e inevitable de nuestra creación. Los números primos, por supuesto, son, de manera semejante, hechos no intencionales, autónomos y objetivos; y en este caso, es obvio que tenemos que *descubrir* otros muchos hechos relativos a ellos: hay conjeturas, como la de Goldbach. Y estas conjeturas, aunque se refieren indirectamente a objetos que son creación nuestra, se refieren directamente a problemas y a hechos que de algún modo surgieron de nuestra creación y sobre los que no podemos tener ni influencia ni control: son hechos crudos, y, a veces, la verdad acerca de ellos es difícil de descubrir.

Esto ejemplifica lo que quiero decir cuando afirmo que el mundo número 3 es en gran parte autónomo, aunque lo hayamos creado los hombres.

Pero tal autonomía es sólo parcial: los nuevos problemas nos llevan a nuevas creaciones o construcciones —tales como las funciones recursivas, o las secuencias de opción libre de Brouwer— y así pueden añadir nuevos objetos al mundo número 3. Y cada uno de estos pasos creará *nuevos hechos no intencionales, nuevos problemas inesperados*, y a menudo, también, *nuevas refutaciones*.⁶

⁶ Un ejemplo de esto último es la "refutación del **estiramiento de los conceptos**" de Lakatos, en I. Lakatos, *Proofs and Refutations [Pruebas y refutaciones]*, 1976, especialmente las pp. 83-99.

Existe también un importantísimo efecto de retroalimentación de nuestra creación sobre nosotros mismos: del mundo número 3 sobre el mundo número 2. Porque los nuevos problemas emergentes nos estimulan a hacer nuevas creaciones.

Podemos describir este proceso con el siguiente esquema, acaso un tanto supersimplificado:⁷

TT ->-EE-

Es decir, empezamos a partir de algún problema p., procedemos a formular alguna solución tentativa o teoría tentativa TT, que puede ser (total o parcialmente) errónea; en cualquier caso, estará sujeta a error-eliminación, EE, que puede consistir en una discusión crítica o en pruebas experimentales; en cualquier caso, surgen nuevos problemas, P₂, de nuestra propia actividad creativa; y estos nuevos problemas, en general, no son intencionalmente creación nuestra, sino que emergen de manera autónoma del campo de nuevas relaciones que no podemos impedir que existan con cada acción nueva, por poco que lo intentemos.

La autonomía del mundo número 3, y la retroalimentación del mundo número 3 sobre el mundo número 2 e incluso sobre el mundo número 1, están entre los hechos más importantes del crecimiento del conocimiento.

Prosiguiendo nuestras consideraciones de índole biológica, resulta fácil advertir que son de importancia general para la teoría de la evolución darviniana: explican cómo podemos levantarnos y andar sin muletas. O, en terminología intelectual, nos ayudan a explicar la "emergencia".

IV. LENGUAJE, CRÍTICA Y EL MUNDO NÚMERO 3

La más importante de las creaciones humanas, con los efectos de retroalimentación más importantes en nosotros mismos, y especialmente en nuestros cerebros, son las funciones superiores del lenguaje humano: más especialmente, la *función descriptiva* y la *función argumentativa*.

Los lenguajes humanos tienen en común con los lenguajes animales las dos funciones inferiores del lenguaje: 1) la autoexpresión y 2) la capacidad de enviar señales. La función autoexpresiva, o función sintomática del lenguaje, es obvia: todo lenguaje animal es sintomático del estado de algún organismo. La función de señalización, de enviar o emi-

⁷ Véase también *Objective Knowledge [Conocimiento objetivo]*, p. 243.

tir señales, también es obvia: no llamamos lingüístico a ningún síntoma a menos que supongamos que puede suscitar una respuesta en otro organismo.

Todos los lenguajes animales y todos los fenómenos lingüísticos comparten estas dos funciones inferiores. Pero el lenguaje humano tiene otras muchas funciones (por ejemplo, la consultiva, la exhortativa y la ficticia o novelesca). Por extraño que parezca, las más importantes de las funciones superiores del lenguaje fueron pasadas por alto por casi todos los filósofos. La explicación de este extraño hecho es que las dos funciones inferiores están siempre presentes cuando están presentes las superiores, de manera que siempre es posible "explicar" cada fenómeno lingüístico, en términos de las funciones inferiores, como una "expresión" o como una "comunicación".

Las dos más importantes funciones de los lenguajes humanos son: 3) la función *descriptiva* y 4) la función *argumentativa*.⁸

Con la función descriptiva del lenguaje humano emerge la idea regulativa de *verdad*, es decir, de una descripción que se apega a los hechos. Otras ideas regulativas o evaluativas son: el contenido, el contenido de verdad y la verosimilitud.⁹

La función argumentativa del lenguaje humano presupone la función descriptiva: los argumentos son, fundamentalmente, acerca de descripciones; critican descripciones desde el punto de vista de las ideas regulativas de verdad, contenido y verosimilitud.

Ahora bien, en esto son importantes dos puntos:

1) Sin el desarrollo de un lenguaje descriptivo exosomático —un lenguaje que, como una herramienta, se desarrolla fuera del cuerpo—, *no puede haber objeto* de nuestra discusión crítica. Pero con el desarrollo de un lenguaje descriptivo propiamente dicho (y, más adelante, de un lenguaje escrito) puede emerger un mundo número 3 lingüístico; y es sólo de esta manera, y sólo en el mundo número 3, como los problemas y las normas del racionalismo crítico pueden desarrollarse.

⁸ Véase *Conjectures and Refutations [Conjeturas y refutaciones]*, capítulos 4 y 12, especialmente las referencias, en las pp. 134, 293 y 295, a K. Bühler, *Sprachtheorie (Teoría del lenguaje)*, 1934. Bühler fue el primero que estudió la decisiva diferencia entre las funciones inferiores y la función descriptiva. Descubrí después, como consecuencia de mi teoría de la crítica, la distinción decisiva entre las funciones descriptiva y argumentativa. Véase también *Objective Knowledge [Conocimiento objetivo]*, pp. 235-238 [y la sección n del escrito 21, más adelante].

⁹ Uno de los grandes descubrimientos de la lógica moderna fue el restablecimiento que hizo Alfred Tarski de la correspondencia (objetiva) entre la teoría de la verdad (verdad = correspondencia con los hechos). Esta teoría, y las ideas reguladoras del contenido de verdad y verosimilitud, se analizan más adelante [en el escrito 14]. El presente ensayo es deudor en todo de la teoría de Tarski; pero, por supuesto, no deseo implicarlo en los crímenes cometidos aquí.

2) Es a este desarrollo de las funciones superiores del lenguaje al que debemos nuestra calidad de seres humanos, nuestra razón. Porque nuestros poderes de razonamiento no son sino poderes de argumentación crítica.

Este segundo punto nos muestra la futilidad de todas las teorías del lenguaje humano que se enfocan en la *expresión* y en la *comunicación*. Como veremos más adelante [en los capítulos 20 y 21], el organismo humano que, como se dice a menudo, tiene la intención de expresarse depende en su estructura, en gran medida, de la emergencia de las dos funciones superiores del lenguaje.

Con la evolución de la función argumentativa del lenguaje, la crítica se convierte en el principal instrumento de un mayor crecimiento. (La lógica puede considerarse *el órgano de la crítica*.)¹⁰ El mundo autónomo de las funciones superiores del lenguaje se convierte en el mundo de la ciencia. Y el esquema, originalmente válido tanto para el mundo animal como para el hombre primitivo,

TT

se convierte en el esquema del crecimiento del conocimiento a través del error-eliminación, por medio de la *crítica racional* sistemática. Se convierte en el esquema de la búsqueda de la verdad y del contenido por medio de la discusión racional. Describe cómo nos levantamos a andar por nuestros propios pies. Nos da una descripción racional de la emergencia evolutiva, y de nuestra *autotrascendencia por medio de la selección y de la crítica racional*.

En resumen, aunque el significado de "conocimiento", como el de todas las palabras, no es importante, sí importa para distinguir entre diferentes sentidos o acepciones de esta palabra:

1) El conocimiento subjetivo, que consiste en ciertas disposiciones innatas para actuar, y en sus modificaciones adquiridas.

2) El conocimiento objetivo; por ejemplo, el conocimiento científico, que consiste en teorías conjeturales, problemas abiertos, situaciones problemáticas y argumentos.

Todo trabajo, en ciencia, es trabajo dirigido hacia el desarrollo del conocimiento objetivo. Somos trabajadores que estamos añadiendo algo al desarrollo del conocimiento objetivo, así como los albañiles trabajan en una catedral.

io Véase *Conjectures and Refutations* [*Conjeturas y refutaciones*], p. 64.

Nuestro trabajo es falible, como todo trabajo humano. Cometemos errores constantemente, y existen estándares y normas de objetividad que acaso no alcancemos a satisfacer: normas de verdad, de contenido, de validez, y otras.

El lenguaje, la formulación de problemas, la emergencia de nuevas situaciones problemáticas, las teorías en conflicto entre ellas mismas, la crítica mutua por medio de la argumentación: todos estos elementos son medios indispensables para lograr el desarrollo científico. Las más importantes funciones o dimensiones del lenguaje humano son la descriptiva y la argumentativa (que no posee el lenguaje animal). El desarrollo de estas funciones constituye, por supuesto, nuestro verdadero trabajo, aunque son consecuencias no intencionales de nuestras acciones. Sólo dentro de un lenguaje así enriquecido se vuelven posibles el argumento de crítica y el conocimiento, en su sentido objetivo.

La repercusión, o los efectos de retroalimentación de la evolución del mundo número 3 en nosotros mismos, en nuestros cerebros, en nuestras tradiciones (si cualquier hombre tuviera que empezar por donde empezó Adán, no llegaría más lejos de donde llegó Adán), en nuestras disposiciones para actuar (esto es, en nuestras creencias),¹¹ y en nuestras acciones, no puede subestimarse.

En tanto que opuesta a todo esto, la *epistemología tradicional* está interesada en el mundo número 2: en el conocimiento como cierta clase de creencias; creencias justificables, como la creencia basada en la percepción. En consecuencia, esta clase de filosofía de creencias no puede explicar (y ni siquiera trata de explicar) el decisivo fenómeno de que los científicos critican sus teorías y así las matan. *Los científicos tratan de eliminar sus teorías falsas; tratan de que mueran estas teorías, en vez de que ellos mueran. El creyente —ya sea animal o ser humano— perece con sus creencias falsas.*

V. OBSERVACIONES DE CARÁCTER HISTÓRICO

(i) *Platón y el neoplatonismo*. Por todo cuanto sabemos, Platón fue el descubridor del mundo número 3. Como señaló Whitehead, toda la filosofía occidental consiste en notas al pie de página acerca de los escritos de Platón.

Sólo haré tres breves observaciones sobre Platón, críticas dos de ellas:

¹¹ La teoría de que nuestras creencias pueden medirse por nuestra disposición a apostar por ellas era bien conocida desde 1781; véase I. Kant, *Crítica de la razón pura*, 2ª ed., 1787, p. 853.

1) Platón descubrió, no sólo el mundo número 3, sino parte de la influencia o de la retroalimentación del mundo número 3 sobre nosotros: se dio cuenta de que tratamos de captar las ideas de su mundo número 3; y también advirtió que las utilizamos como explicaciones.

2) El mundo número 3 de Platón era divino; era inmodificable y, por supuesto, verdadero. Así, hay una gran brecha entre su mundo número 3 y mi mundo número 3: mi mundo número 3 es hechura del hombre, y cambiabile, mudable. Contiene, no sólo teorías verdaderas, sino también falsas, y especialmente problemas abiertos, conjeturas y refutaciones.

Y mientras que Platón, el gran maestro de la refutación dialéctica, vio en estos argumentos sólo un medio para llegar al mundo número 3, yo considero a los argumentos los habitantes de mayor importancia del mundo número 3; por no hablar de los problemas abiertos.

3) Platón creía que el mundo número 3, el mundo de las Formas o de las Ideas, nos daría las explicaciones definitivas (esto es, explicaciones por esencias [véanse, más adelante, las pp. 181-182]). Así, por ejemplo, escribe (en *Fedón*, 100 C): "Creo que si cualquier cosa, aparte de la idea de la belleza absoluta, es bella, entonces es bella *por la única razón* de que comparte algo de la idea de belleza absoluta. *Y esta clase de explicación se aplica a todo.*"

Ésta es la teoría de la explicación *fundamental*; es decir, de una explicación cuyo explicando no es capaz de más explicación, ni la necesita. Y es una teoría de *explicación por las esencias*, es decir, por palabras convertidas en hipóstasis.

En consecuencia, Platón consideraba a los objetos del mundo número 3 como algo parecido a cosas inmateriales, o quizá como estrellas o constelaciones: algo para contemplarse, para intuir, pero que no podía tocarse con nuestras mentes. Por esta razón, las habitantes del mundo número 3 de Platón —las Formas o las Ideas— se convirtieron en conceptos de cosas, o esencias, o naturalezas de cosas, y no en teorías, ni en argumentos, ni en problemas.

Esto tuvo las consecuencias de mayor alcance para la historia de la filosofía. Desde Platón hasta nuestros días, la mayoría de los filósofos han sido, o nominalistas,¹² o lo que he llamado esencialistas. Les interesa más el significado (esencial) de las palabras que la verdad o la falsedad de sus teorías.

Presento a menudo el problema en forma de cuadro (página siguiente). Mi tesis es que *el lado izquierdo de este cuadro no tiene importancia*,

¹² Véase J. W. N. Watkins, *Hobbes's System of Ideas [El sistema de ideas de Hobbes]*, 1965, capítulo vin, especialmente las pp. 145 y ss; *The Logic of Scientific Discovery [La lógica de la investigación científica]*, pp. 420-422, y *Conjectures and Refutations [Conjeturas y refutaciones]*, pp. 18 y ss, 262 y 297 y ss.

IDEAS

es decir,

DESIGNACIONES O TÉRMINOS
O CONCEPTOS

DECLARACIONES O PROPOSICIONES
O TEORÍAS

pueden formularse con

PALABRAS

AFIRMACIONES

que pueden ser

SIGNIFICANTES

VERDADERAS

y su

SIGNIFICADO

VERDAD

puede reducirse, mediante

DEFINICIONES

DERIVACIONES

CONCEPTOS NO DEFINIDOS

PROPOSICIONES PRIMITIVAS

Intentar establecer (en vez de reducir) por estos medios su

SIGNIFICADO

VERDAD

nos lleva a una regresión infinita

comparado con el lado derecho: lo que debería interesarnos son las teorías, la verdad, los argumentos. Si tantos filósofos y científicos piensan todavía que los conceptos y los sistemas conceptuales (y sus problemas de significado, o el significado de las palabras) son comparables en importancia a las teorías y a los sistemas teóricos (y a los problemas de su veracidad, o al de la verdad de las declaraciones), entonces están sufriendo el mismo error principal de Platón.¹³ Porque los conceptos son,

¹³ El error, que es tradicional, se conoce como "el problema de los universales". Esto debería sustituirse con "el problema de las teorías", o con "el problema del contenido teóri-

en parte, medios para formular teorías, y en parte, medios para resumir teorías. En cualquier caso, su significación es primordialmente instrumental, y siempre pueden sustituirse por otros conceptos.

Los contenidos y objetos del pensamiento parecen haber desempeñado un papel importante en el estoicismo y en el neoplatonismo. Plotino conservó la separación que hizo Platón entre el mundo empírico y el mundo de las Formas y las Ideas. Sin embargo, como Aristóteles,¹⁴ Plotino destruyó la trascendencia del mundo de Platón, al colocarla en la conciencia de Dios.

Plotino criticó a Aristóteles por no haber podido distinguir entre la Primera Hipóstasis (la Unicidad) y la Segunda Hipóstasis (el intelecto divino). Pero siguió a Aristóteles al identificar los actos de pensamiento de Dios con sus propios contenidos u objetos; y elaboró este punto de vista al considerar que las Formas o Ideas del mundo inteligible de Platón son los estados de conciencia inmanentes del intelecto divino.¹⁵

(ii) *Hegel*. Hegel era un platónico (o, más bien, una especie de neoplatónico) y, como Platón, una especie de heraclitano. Era un platónico cuyo mundo de ideas estaba cambiando, evolucionando. Las "Formas" o "Ideas" de Platón eran objetivas, y no tenían nada que ver con las ideas conscientes en una mente subjetiva; habitaban en un mundo divino, inmutable, celestial (supralunar, según Aristóteles). En contraste, las ideas de Hegel, como las de Plotino, eran fenómenos conscientes: los pensamientos se pensaban a sí mismos y habitaban en una especie de conciencia, en una especie de mente o "Espíritu"; y junto con este "Espíritu", estaban cambiando o evolucionando. El hecho de que el "Espíritu Objetivo" y el "Espíritu Absoluto" de Hegel estén sujetos a cambio es el único punto en que sus Espíritus son más semejantes a mi "mundo

co de todo el lenguaje humano". Véase *La lógica de la investigación científica*, nota *1 a la sección IV [y también la sección I del escrito 11, más adelante].

Incidentalmente, está claro que, de las tres posiciones —*universale ante rem*, *in re*, y *post rem*—, la última, en su significado usual, es anti-mundo número 3 y trata de explicar el lenguaje como expresión, mientras que la primera (platónica) es pro-mundo número 3. Resulta interesante que la posición intermedia (aristotélica) (*in re*) puede considerarse, o bien anti-mundo número 3, o que pasa por alto el problema del mundo número 3. De esta manera, da testimonio de la influencia distorsionadora del conceptualismo.

¹⁴ Véase Aristóteles, *Metafísica*, 1072b21 y ss y 1074b15-1075a4. Este pasaje (que Ross resume así: "el pensamiento divino debe ocuparse del objeto más divino, que es él mismo") contiene una crítica implícita de Platón. Su afinidad con las ideas platónicas está especialmente clara en las líneas 25 y ss: "piensa en lo que es más divino y no cambia; porque el cambio sería cambio para lo peor..." (Véase Aristóteles, *De Anima*, 429b27 y ss, especialmente 430a4.)

¹⁵ Véase Plotino, *Ennéadas* II.iv.4, III.viii.11, V.iii.2-5, V.ix.5-8, VI.v.2 y VI.vi.6-7. (La numeración que aquí se reproduce es la de la edición de R. Volkmann, 1883; hay ligeras variaciones en la traducción inglesa de S. Mackenna, 1917-1930.)

número 3" que el mundo de Ideas de Platón (o el de Bolzano, el mundo de las "afirmaciones como tales").

Las diferencias más importantes entre el "Espíritu Objetivo" y el "Espíritu Absoluto" de Hegel y mi "mundo número 3" son las siguientes:

1) Según Hegel, aunque el Espíritu Objetivo (que abarca a la creación artística) y el Espíritu Absoluto (que abarca a la filosofía) consisten en producciones humanas, el hombre no es creativo. Es el Espíritu Objetivo convertido en hipóstasis, es la divina autoconciencia del Universo, lo que mueve al hombre: "los individuos... son instrumentos", instrumentos del Espíritu de la Época, y su trabajo, su "ocupación sustancial", está "preparada y señalada independientemente de ellos".¹⁶

Así pues, lo que he llamado la autonomía del mundo número 3, y su efecto de retroalimentación, se vuelven omnipotentes para Hegel; tal es sólo uno de los aspectos de su sistema en que se manifiesta su formación teológica. En contra de esto, asevero que el elemento creativo individual, la relación de toma y daca entre un hombre y su trabajo, reviste la mayor importancia. En Hegel, esto degenera hasta convertirse en la doctrina de que el gran hombre es una especie de médium en el que se manifiesta el Espíritu de la Época.

2) Pese a cierta similitud superficial entre la dialéctica de Hegel y mi esquema evolutivo

- • TT

existe una diferencia fundamental. Mi esquema funciona mediante error-eliminación y, en el nivel científico, mediante la crítica consciente, basado en la idea reguladora de la búsqueda de la verdad.

La crítica consiste, por supuesto, en la búsqueda de las contradicciones y en su eliminación: la dificultad creada por la exigencia de su eliminación constituye el nuevo problema (P_2). Así, la eliminación del error conduce al desarrollo objetivo de nuestro conocimiento; del conocimiento en su acepción objetiva. Conduce también al desarrollo de la verosimilitud objetiva: posibilita nuestra aproximación a la verdad (absoluta).

Por otra parte, Hegel es un relativista.¹⁷ No ve nuestra tarea como la búsqueda de las contradicciones, con la meta de eliminarlas, porque piensa que las contradicciones son tan buenas como los sistemas teóricos no

¹⁶ Véase G. W. F. Hegel, *Enzyklopadie der Philosophischen Wissenschaften*, 3ª ed., 1930, párrafo 551. (Hay traducción al inglés por W. Wallace, disponible en A. V. Miller (comp.), *Hegel's Philosophy of Mind [Filosofía de la mente de Hegel]*, 1971.)

¹⁷ Véase *Conjeturas y refutaciones*, capítulo 15, y *La sociedad abierta y sus enemigos*, apéndice I ("Hechos, normas y verdad: más críticas al relativismo") al volumen II.

contradictorios (o mejores): nos proveen del mecanismo mediante el cual el Espíritu se impulsa a sí mismo. Así pues, la crítica racional no desempeña ningún papel en el automatismo hegeliano, y tampoco desempeña en él ningún papel la creatividad humana.¹⁸

3) Mientras que Platón deja que sus Ideas convertidas en hipóstasis habiten en algún cielo divino, Hegel personaliza a su Espíritu en alguna conciencia humana. Su doctrina consiste, cabalmente, en que el Espíritu no sólo es consciente, sino una persona, un yo. Contra esto hay que decir que mi mundo número 3 no tiene similitud alguna con la conciencia humana; y aunque sus primeros habitantes son productos de la conciencia humana, son del todo diferentes de las ideas conscientes o de los pensamientos en la acepción subjetiva de esta palabra.

5. EPISTEMOLOGÍA EVOLUTIVA (1973)

AHORA me ocuparé del progreso en las ciencias. Estudiaré el progreso científico desde el punto de vista biológico o evolutivo. Estoy lejos de sugerir que éste es el punto de vista más importante para examinar los avances de la ciencia, pero el enfoque biológico me da un medio conveniente para introducir las dos ideas rectoras de la primera mitad de mi conferencia. Estas dos ideas son la *instrucción* y la *selección*.

Desde el punto de vista biológico o evolutivo, la ciencia, o el avance científico, puede considerarse una forma que utiliza la especie humana para adaptarse al medio ecológico: para invadir nuevos nichos ecológicos, e incluso para inventar nuevos nichos ecológicos.¹ Esto nos lleva a exponer el siguiente problema.

Podemos distinguir tres niveles de adaptación: la adaptación genética, el aprendizaje conductual adaptativo y el descubrimiento científico, que es un caso especial del aprendizaje conductual adaptativo. El problema principal que me ocupa en esta parte de mi conferencia consiste en indagar las semejanzas y desemejanzas entre las estrategias de progreso o de adaptación en el nivel *científico* y en esos otros dos niveles: el *genético* y el *conductual*. Y compararé los tres niveles de adaptación investigando el papel que desempeñan en cada uno de ellos la *instrucción* y la *selección*.

Para no dejar a mis oyentes y al lector con una venda en los ojos en cuanto al resultado de esta comparación, anticiparé de inmediato mi tesis principal. Es una tesis que afirma la *similitud fundamental de los tres niveles*, como sigue.

En los tres niveles —los de adaptación genética, conducta adaptativa y descubrimiento científico—, el mecanismo de adaptación es fundamentalmente el mismo.

¹ La formación de proteínas con membrana, de los primeros virus, y de células, quizás haya sido una de las primeras invenciones de los nuevos nichos ambientales; aunque es posible que otros nichos ambientales (quizá redes de enzimas inventadas por otros genes desnudos de otras formas) se hayan inventado más temprano.

¹⁸ Véase Lakatos, *op. cit.*, nota 2 en la p. 54.

Esto puede explicarse con cierto detalle.

La adaptación, en los tres niveles, empieza a partir de una *estructura* heredada que es básica. En el nivel genético es la *estructura de genes del organismo*. A esta estructura corresponde, en el nivel conductual, el *repertorio innato* de los tipos de conducta de que dispone el organismo; y en el nivel científico, las *conjeturas o teorías científicas dominantes*. Estas *estructuras* se transmiten siempre por *instrucción* en los tres niveles: por replicación de la instrucción genética codificada, en los niveles genético y conductual; y mediante tradición social e imitación, en los niveles conductual y científico. En los tres niveles, la *instrucción* procede desde *dentro de la estructura*. Si ocurren mutaciones, variaciones o errores, entonces éstas son nuevas instrucciones, que también surgen *desde el interior de la estructura*, y no desde afuera, no desde el medio biológico.

Estas estructuras heredadas están expuestas a ciertas presiones, o retos, o problemas: a presiones de selección, a retos ambientales, a problemas teóricos. Como respuesta, se originan variaciones de las *instrucciones* genética o tradicionalmente heredadas,² con métodos que son, por lo menos parcialmente, *fortuitos*. En el nivel genético, éstas son mutaciones y recombinaciones de la instrucción codificada; en el nivel conductual, son variaciones provisorias y recombinaciones en el repertorio; en el nivel científico, son nuevas y revolucionarias teorías provisionales. En los tres niveles, obtenemos nuevas instrucciones tentativas, o, en una palabra, tanteos.

Es importante señalar que estos intentos tentativos son cambios que se originan *dentro* de la estructura individual de manera más o menos fortuita... en los tres niveles. El punto de vista de que *no* se deben a instrucciones procedentes del exterior, del entorno, es apoyado (aunque sólo débilmente) por el hecho de que muy diversos organismos pueden responder a veces, de maneras muy distintas, al mismo nuevo reto ambiental.

La siguiente etapa es la de *selección* a partir de las mutaciones y variaciones disponibles: las de nuevos intentos o tanteos que están mal adaptadas son eliminadas. *Ésta es la etapa de la eliminación del error*. Sólo las instrucciones tentativas más o menos bien adaptadas sobreviven y, a su vez, son heredadas. Así pues, podemos hablar de *adaptación por el "método de prueba y error"*, o, mejor dicho, por "el método de

prueba y eliminación del error". La eliminación del error, o de las instrucciones mal adaptadas, se llama también "*selección natural*": es una especie de "retroalimentación negativa", y opera en los tres niveles.

Cabe señalar que, en general, no se llega a *ningún estado de equilibrio de adaptación* mediante ninguna aplicación del método de prueba y eliminación del error, ni mediante la selección natural. En primer lugar, porque no es factible que se presenten soluciones tentativas perfectas u óptimas para resolver el problema; y en segundo lugar —y esto es más importante—, porque la emergencia de estructuras nuevas, o nuevas instrucciones, implica un cambio en la situación del medio biológico. Acaso se vuelvan importantes nuevos elementos ambientales: por consiguiente, pueden surgir nuevas presiones, nuevos cambios, nuevos problemas, resultado de los cambios estructurales que se han presentado en el interior del organismo.

En el nivel genético, el cambio puede ser la mutación de un gene, con el consiguiente cambio de una enzima. Ahora bien, la red de enzimas forma el medio biológico más íntimo de la estructura genética. Por tanto, habrá un cambio en este medio biológico íntimo; y con él, pueden suscitarse nuevas relaciones entre el organismo y el más remoto medio biológico, exterior al organismo; y habrá también nuevas presiones de selección.

En el nivel conductual sucede lo mismo; porque la adopción de una nueva clase de comportamiento puede equivaler, en la mayoría de los casos, a la adopción de un nuevo nicho ecológico. Por consiguiente, surgirán nuevas presiones de selección, y habrá nuevos cambios genéticos.

En el nivel científico, la adopción tentativa de una nueva conjetura o teoría puede resolver uno o dos problemas, pero invariablemente abre la puerta a muchos *nuevos* problemas; porque una nueva teoría revolucionaria funciona exactamente como un nuevo y poderoso órgano de significación. Si el progreso es importante, los nuevos problemas diferirán de los viejos problemas: los nuevos problemas estarán en un nivel de profundidad radicalmente distinto. Tal fue lo que ocurrió, por ejemplo, con la teoría de la relatividad; y también ocurrió esto con la mecánica cuántica; y es lo que está ocurriendo ahora mismo con la biología molecular. En cada uno de estos casos, la nueva teoría abrió nuevos horizontes de problemas inesperados.

Sugiero que ésta es la manera como progresa la ciencia. Y nuestro progreso puede medirse comparando nuestros viejos problemas con los nuevos que se nos van presentando. Si el progreso que se ha logrado es grande, los nuevos problemas serán de carácter insospechado. Habrá problemas de mayor profundidad; y, además, los habrá en mayor número. Cuanto más avancemos en conocimiento, advertiremos más cómo

² Constituye un problema abierto a discusión si podemos hablar en estos términos ("en respuesta") acerca del nivel genético (compárese con mi conjetura acerca de los mutágenos respondientes, en la sección v). Sin embargo, si no hubiese variaciones, no podría haber ni adaptación ni evolución; por tanto, podemos decir que la presentación de mutaciones es parcialmente controlada por una necesidad de que las haya, o funciona como si hubiera tal necesidad.

ramente la vastedad de nuestra ignorancia. (La toma de conciencia de nuestra ignorancia se ha agudizado, por ejemplo, como resultado de la asombrosa revolución que suscitó hace poco la biología molecular.)

A continuación, resumiré mi tesis.

En los tres niveles que estoy considerando: el genético, el conductual y el científico, estamos operando con estructuras heredadas que se transmiten mediante instrucción; ya sea mediante el código genético o a través de la tradición. En estos tres niveles surgen nuevas estructuras y nuevas instrucciones de los cambios originados *en el interior de la estructura*: por pruebas tentativas que están sujetas a la selección natural o a la eliminación del error.

ni

Hasta ahora he recalcado las *semejanzas* en el funcionamiento del mecanismo adaptativo en los tres niveles. Esto nos plantea un problema obvio: ¿Qué decir de las *diferencias*?

La diferencia principal entre los niveles genético y conductual es ésta: las mutaciones en el nivel genético son, no sólo azarosas, sino del todo "ciegas", en dos sentidos. En primer lugar, de ninguna manera están orientadas hacia metas por alcanzar. En segundo lugar, la supervivencia de una mutación no puede ejercer influencia en otras mutaciones posteriores, ni siquiera en las frecuencias o en las probabilidades de que se presenten; aunque sea cierto que la *supervivencia* de una mutación a veces puede determinar qué clase de mutaciones posiblemente *sobrevivan* en casos futuros. En el nivel conductual, los intentos son también más o menos fortuitos, pero no son ya completamente "ciegos", en ninguno de los dos sentidos que hemos mencionado. Porque, primero, están dirigidos hacia metas por alcanzar; segundo, los animales pueden aprender del resultado de un intento: pueden aprender a evitar el tipo de conducta tentativa que los ha llevado a fracasar. (Pueden, incluso, evitarlo en casos en que posiblemente habría resultado un éxito.) De manera semejante, también pueden aprender del éxito; y la conducta fructífera puede repetirse, aun en casos en que no sea la apropiada a las circunstancias. Sin embargo, cierto grado de "ceguera" es inherente a todos los intentos.³

³ En cuanto a la utilización del término "ciego" (especialmente en la segunda acepción), véase D. T. Campbell, "Methodological Suggestions for a Comparative Psychology of Knowledge Processes" ["Sugerencias metodológicas para la psicología comparativa de los procesos del conocimiento"], en *Inquiry* 2, 1959, pp. 152-182; "Blind Variation and Selective Retention in Creative Thought as in Other Knowledge Processes" ["Variación ciega y retención selectiva en el pensamiento creativo y en otros procesos del conocimiento"], en *Psychological Review* 67, 1960, pp. 380-400; y "Evolutionary Epistemology" ["Epis-

La adaptación conductual es generalmente un proceso activo e intensivo: el animal —especialmente el animal joven cuando juega— e incluso la planta están investigando activamente el medio biológico.⁴

Esta actividad, que está genéticamente programada en gran medida, en mi opinión marca una importante diferencia entre el nivel genético y el nivel conductual. En esto, puede compararse con la experiencia que los psicólogos de la *Gestalt* llaman "*insight*" ["perspicacia"]: es una experiencia que acompaña a muchos descubrimientos conductuales.⁵ No obstante, no debe pasarse por alto que hasta un descubrimiento acompañado de "perspicacia" puede ser *erróneo*; cada intento, incluso uno con "perspicacia", es de naturaleza conjetural o hipotética. Recordemos que los monos de Köhler a veces dan con "perspicacia" en lo que a la postre resulta un intento fallido de resolver su problema; y hasta a los grandes matemáticos suele desorientarlos una intuición. Así pues, tanto los animales como los hombres tienen que poner a prueba sus hipótesis; tienen que utilizar el método de prueba y eliminación de error.

temología evolutiva"], pp. 413-463 de P. A. Schilpp (comp.), *The Philosophy of Karl Popper*, The Library of Living Philosophers [Biblioteca de Filósofos Vivientes], 1974.

Mientras que la "ceguera" de las pruebas se relaciona con lo que hemos descubierto en el pasado, el azar se relaciona con un conjunto de elementos (que forman la "muestra de espacio"). En el nivel genético, estos "elementos" son las cuatro bases de nucleoproteínas; en el nivel conductual, son los constituyentes del repertorio de conductas del organismo. Estos constituyentes pueden tener diferentes pesos respecto a diferentes necesidades u objetivos, pesos que pueden cambiar a través de la experiencia (bajando el grado de "ceguera").

⁴ Sobre la importancia de la participación activa, véase R. Held y A. Hein, "Movement-produced Stimulation in the Development of Visually Guided Behaviour" ["Estimulación de producción de movimiento en el desarrollo de conducta guiada visualmente"], en *Journal of Comparative and Physiological Psychology* 56, 1963, pp. 872-876; véase también J. C. Eccles, *Facing reality* [Enfrentándonos a la realidad], 1963, pp. 66 y ss. La actividad es, al menos parcialmente, una actividad que produce hipótesis: véase J. Krechevsky, "Hypothesis' versus 'Chance' in the Pre-solution Period in Sensory Discrimination-learning" ["Hipótesis contra azar en el periodo de presolución en discriminación-aprendizaje sensorial"], en *University of California Publications in Psychology* 6, 1932, pp. 27-44 (reproducido en las pp. 183-197 de A. J. Riopelle (comp.), *Animal Problem Solving* [Resolución de problemas en los animales], 1967).

⁵ Quizá convenga ocuparme aquí de algunas de las diferencias entre mis puntos de vista y los puntos de vista de la escuela de la *Gestalt*. (Por supuesto, acepto el hecho de la percepción *Gestalt*; sólo tengo dudas sobre lo que podríamos llamar la filosofía *Gestalt*.)

Conjeturo que la unidad, o la articulación, de la percepción depende más estrechamente de los sistemas de control motor y de los sistemas neurales eferentes del cerebro, que de los sistemas aferentes; que esta unidad de la percepción depende estrechamente del repertorio conductual del organismo. También conjeturo que una araña o un ratón nunca tendrán una intuición (como la tenía el simio de Köhler) sobre la posible unidad de las dos varas que pueden juntarse, porque manejar varas de ese tamaño no forma parte de su repertorio conductual. Todo esto puede interpretarse como una especie de generalización de la teoría de las emociones de James/Lange (1884; véanse las pp. 449 y ss de W. James, *The Principles of Psychology* [Principios de psicología], tomo II, 1890), que hace extensiva la teoría de nuestras emociones a nuestras percepciones (especialmente de las

Por otra parte, estoy de acuerdo con Köhler y con Thorpe⁶ en que los intentos que hacen los animales en la resolución de problemas en general no son completamente "ciegos". Sólo en casos extremos, cuando el problema al que se enfrenta el animal no concuerda con la formulación de hipótesis, el animal recurre a intentos más o menos ciegos y fortuitos para salir de una situación desconcertante. Sin embargo, aun en estos intentos, generalmente es discernible la orientación hacia el logro de metas, en gran contraste con lo fortuito y ciego de las mutaciones y recombinaciones genéticas.

Otra diferencia entre los cambios genéticos y los cambios conductuales adaptativos es que los primeros *siempre* establecen una estructura genética rígida y casi invariable. Los segundos, según se admite, también conducen *a veces* a una pauta de conducta rígida, a la que se adhiere dogmáticamente el sujeto; radicalmente, en el caso de las "marcas indelebles" [*imprinting*] (según Konrad Lorenz); pero en otros casos llevan al sujeto a adoptar una pauta conductual flexible, que permite la diferenciación o la modificación; por ejemplo, puede suscitar una conducta exploratoria, o lo que Pavlov llamó el "reflejo de libertad".⁷

En el nivel científico, los descubrimientos son revolucionarios y creativos. No cabe duda de que debe atribuirse cierta creatividad a todos los niveles, incluso al genético: nuevos ensayos, que originan nuevos ambientes, y por ende, nuevas presiones de selección, crean nuevos y revolucionarios resultados en todos los niveles, aunque hay marcadas tendencias conservadoras inducidas en los diferentes mecanismos de instrucción.

percepciones de tipo *Gestalt*), que de esta manera no se nos "darían" (como en la teoría de la *Gestalt*), sino que nosotros las "haríamos" al decodificar claves ("dadas" comparativamente). El hecho de que las claves pueden desorientar (ilusiones ópticas, en el hombre, ilusiones con fantoches o muñecos de trapo, en los animales, etc.) puede explicarse por la necesidad biológica de imponer nuestras interpretaciones conductuales en claves muy simplificadas. La conjetura de que nuestra decodificación de lo que nos dicen nuestros sentidos depende de nuestro repertorio conductual puede explicar parte de la gran diferencia que existe entre los animales y el hombre; porque, a través de la evolución del lenguaje humano, nuestro repertorio ha llegado a ser ilimitado.

⁶ Véanse las pp. 99 y ss de W. H. Thorpe, *Learning and Instinct in Animals* [*El aprendizaje y el instinto en los animales*], 1956; edición de 1963, pp. 100-147; y W. Köhler, *The Mentality of Apes* [*La mentalidad de los monos*], 1925; edición Penguin Books, 1957, pp. 166 y ss.

⁷ Véase I. P. Pavlov, *Conditioned Reflexes* [*Reflejos condicionados*], 1927, especialmente las pp. 11 y ss. A la luz de lo que Pavlov llama "conducta exploratoria", y su muy afín "conducta de libertad" —ambas basadas, obviamente, en lo genético—, y en vista de la importancia de estas dos conductas para la actividad científica, me parece que la conducta de los conductistas que apuntan a sustituir el valor de la libertad con lo que ellos llaman "refuerzo positivo" puede ser síntoma de hostilidad inconsciente hacia la ciencia. Incidentalmente, lo que B. F. Skinner, en *Beyond Freedom and Dignity* [*Más allá de la libertad y de la dignidad*], 1972, llama "la literatura de la libertad" no surgió como resultado de un refuerzo negativo, como él sugiere. Más bien surgió, desde Esquilo y Píndaro, como resultado de las victorias de Maratón y Salamina.

La adaptación genética puede operar, por supuesto, sólo dentro del lapso de tiempo de unas cuantas generaciones; cuando mucho, por ejemplo, en el transcurso de una o dos generaciones. En organismos que replican sus genes muy rápidamente, esto puede significar un lapso muy breve; y quizá no haya allí oportunidad para la adaptación conductual. Los organismos de reproducción lenta se ven obligados a inventar la adaptación conductual, para ajustarse a los cambios rápidos del ambiente. Por ello, necesitan disponer de un repertorio conductual, con tipos de comportamiento de mayor duración o de mayor o menor alcance. El repertorio conductual, y la amplitud de los tipos de conducta disponibles, pueden considerarse genéticamente programados; de ahí que, como ya lo hemos indicado, pueda afirmarse que un nuevo tipo de conducta implicará la elección de un nuevo nicho ecológico, y que nuevos tipos de conducta pueden ser, por cierto, genéticamente creativos, pues a su vez pueden determinar nuevas presiones de selección y, con ello, decidir indirectamente la futura evolución de la estructura genética.⁸

En el nivel del descubrimiento científico emergen dos nuevos puntos de vista. El más importante consiste en que las teorías científicas pueden formularse lingüísticamente, y que hasta pueden publicarse. Así se vuelven objetos fuera de nosotros mismos: objetos abiertos a la investigación. En consecuencia, ya están abiertos a la *crítica*. Así podemos descartar una teoría defectuosa antes de que la adopción de esa teoría nos haga ineptos para sobrevivir: al criticar nuestras teorías, podemos hacer que ellas mueran, en vez de que muramos nosotros. Esto, claro está, reviste muchísima importancia.

El otro punto también se relaciona con el lenguaje. Una de las novedades del lenguaje humano es que alienta la narración, y así, *la imaginación creadora*. El descubrimiento científico es afín a la narración explicativa

⁸ Así, la conducta exploratoria y la resolución de problemas crean nuevas condiciones para la evolución de los sistemas genéticos; condiciones que afectan profundamente a la selección natural de estos sistemas. Podríamos decir que en cuanto se ha alcanzado cierta latitud —incluso la alcanzada entre los organismos unicelulares (véase, especialmente, la obra clásica de H. S. Jennings, *The Behaviour of the Lower Organisms* [*La conducta de los organismos inferiores*], 1906)— la iniciativa del organismo al elegir su ecología o hábitat es el factor más importante, y que la selección natural dentro del nuevo hábitat es el segundo factor en importancia. De esta manera, el darwinismo puede simular el lamarckismo, e incluso la "evolución creadora" de Bergson. Esto lo han reconocido darvinistas estrictos. Para una brillante presentación y reseña de la historia, véase Alister Hardy, *The Living Stream* [*La corriente de la vida*], 1965, especialmente las conferencias vi, vil y vm, donde se encontrarán muchas referencias a literatura precedente al respecto, desde James Hutton (muerto en 1797) en adelante (véanse las pp. 178 y ss). Véase también Ernst Mayr, *Animal Species and Evolution* [*Las especies animales y la evolución*], 1963, pp. 604 y ss y p. 611; Erwin Schrödinger, *Mind and Matter* [*Mente y materia*], 1958, capítulo 2; F. W. Braestrup, "The Evolutionary Significance of Learning" ["El significado evolutivo del aprendizaje"], en *Videnskabelige Meddelelser Naturhistorisk Forening i Kjøbenhavn* 134, 1971, pp. 89-102 (con bibliografía), y *Conocimiento objetivo*, capítulo 7.

de cuentos, a la creación de mitos y a la imaginación poética en acción. Y, por supuesto, el desarrollo de la imaginación refuerza la necesidad de tener algún control sobre ella, como sucede, en materia de ciencia, con la crítica interpersonal: con la amistosa-hostil colaboración de los científicos, que en parte se basa en la competición y en parte en la meta común de acercarnos más a la verdad. Esto, y el papel que desempeñan la instrucción y la tradición, me parece que completa los principales elementos sociológicos implicados en el avance de la ciencia; aunque, claro está, podríamos hablar mucho acerca de los obstáculos sociales que se oponen al progreso, o acerca de los peligros sociales inherentes al progreso.

IV

He sugerido que el progreso en la ciencia, o descubrimiento científico, depende de *la enseñanza y de la selección*: de un elemento conservador, tradicional o histórico, y del empleo revolucionario de la prueba y la eliminación del error mediante la crítica, lo cual incluye rigurosos exámenes o pruebas; es decir, que hay que intentar examinar las probables debilidades de las teorías; hay que intentar refutarlas.

Claro está que el científico individual deseará asentar bien su teoría, más que refutarla. Pero, desde el punto de vista del progreso de la ciencia, este deseo puede hacerlo fallar fácilmente. Además, si él mismo no examina críticamente su teoría favorita, otros lo harán por él. Los únicos resultados que los críticos considerarán apoyos de esa teoría serán las fallas de los interesantes intentos de refutarla; fallas en encontrar contraejemplos, allí donde más se esperaría encontrar contraejemplos, a la luz de las mejores teorías que compitan al respecto. Por tanto, no hay necesidad de crear un gran obstáculo para la ciencia si el científico tiene preferencia por una teoría en especial. Pero creo que Claude Bernard fue muy sabio al escribir: "Quienes tienen una fe excesiva en sus ideas, no están bien preparados para hacer descubrimientos."⁹

Todo esto es parte del enfoque crítico de la ciencia, opuesto al enfoque inductivista; o del enfoque darvinista o "eliminacionista" o "seleccionista", opuesto al enfoque de Lamarck, que opera basado en la idea de *la enseñanza desde afuera*, o sea, desde el medio exterior al organismo, mientras que el enfoque "seleccionista" o crítico sólo permite *la enseñanza desde dentro*, desde el interior de la estructura misma.

De hecho, sostengo que *no existe tal cosa como la enseñanza desde el*

⁹ Citado por Jacques Hadamard, *The Psychology of Invention in the Mathematical Field* [La psicología de la invención en el campo de las matemáticas] 1945- edición de Dover 1954, p. 48.

exterior de la estructura, o la pasiva recepción de un flujo de información que se imprima indeleblemente en nuestros órganos sensoriales. Todas las observaciones están impregnadas de teoría: no existe la observación pura, desinteresada, libre de teoría. (Para darnos cuenta de esto podemos intentar, con un poco de imaginación, comparar la observación de un humano con la de una hormiga o la de una araña.)

Francis Bacon estaba en lo justo cuando le preocupaba que nuestras teorías pudieran perjudicar nuestras observaciones. Esto lo indujo a aconsejar a los científicos que evitaran el prejuicio purificando su mente de toda teoría. Todavía en nuestro tiempo se dan recetas de esta índole.¹⁰ Pero, para lograr la objetividad, no podemos confiar en la mente en blanco: la objetividad se apoya en la crítica, en la discusión crítica, y en el examen crítico de los experimentos. [Véanse más adelante el escrito 11, sección n, y el escrito 30.] Y debemos reconocer, especialmente, que nuestros órganos sensoriales mismos incorporan lo que equivale a los prejuicios. Ya he recalado antes [en la sección n] que las teorías son como órganos sensoriales. Ahora deseo recalcar que nuestros órganos sensoriales son como teorías. Ellos *incorporan* teorías adaptativas (como se ha demostrado en el caso de conejos y gatos). Y estas teorías son resultado de la selección natural.

Sin embargo, ni siquiera Darwin o Wallace, por no mencionar a Spencer, advirtieron que no existe la enseñanza desde afuera. Ellos no manejaron argumentos puramente seleccionistas. En realidad, frecuentemente argumentaron siguiendo la línea de Lamarck.*¹ En esto, parece que se equivocaron. Sin embargo, acaso valga la pena especular sobre los probables

¹⁰ Los psicólogos conductistas que estudian "la distorsión del experimentador" han descubierto que algunas ratas albinas se desempeñan decididamente mejor que otras si el experimentador cree (erróneamente) que las primeras pertenecen a una estirpe seleccionada por su inteligencia superior. Véase R. Rosenthal y K. L. Fode, "The Effect of Experimenter Bias on the Performance of the Albino Rat" ["El efecto de la distorsión del experimentador en el desempeño de la rata albina"], en *Behavioural Science* 8, 1963, pp. 183-189. La lección que obtuvieron los autores de este trabajo es que los experimentos deben hacerse "por ayudantes de investigación que no sepan de antemano qué resultado se desea" (p. 188). Como Bacon, estos autores cifran su esperanza en la mente vacía, olvidando que las expectativas del director de la investigación pueden comunicarse, sin indicios explícitos, a sus ayudantes, así como parecen haberse comunicado de cada ayudante de investigación a sus ratas.

¹¹ Es interesante notar que Charles Darwin, en sus últimos años, creía en la ocasional herencia incluso de las mutilaciones. Véase su obra *The Variation of Animals and Plants under Domestication* [La variación de animales y plantas en condiciones de domesticación], tomo i, 2ª ed., 1875, pp. 466-470.

límites del darvinismo; porque debemos estar siempre alerta para buscar opciones que sean diferentes de cualquier teoría dominante.

Pienso que en este asunto debemos aclarar dos puntos: el primero es que el argumento contra la herencia genética de características adquiridas (tales como las mutilaciones) depende de la existencia de un mecanismo genético en el que hay una distinción bastante clara entre la estructura de los genes y el resto del organismo: el soma. Pero este mecanismo genético debe ser, a su vez, un producto tardío de la evolución, y sin duda lo precedieron otros varios mecanismos menos refinados. Ciertas clases muy especiales de mutilaciones *son* realmente heredadas: más particularmente, las mutilaciones por radiación de la estructura de los genes. Así, por ejemplo, si suponemos que el organismo primigenio era un solo gene desnudo, podemos afirmar, incluso, que cada mutilación no letal ocurrida a este organismo sería heredada. Lo que no podemos decir es que este hecho contribuya en alguna forma a explicar la adaptación genética, o el aprendizaje genético, excepto indirectamente, a través de la selección natural.

El segundo punto es el siguiente: podemos considerar la muy tentativa conjetura de que, como respuesta somática a ciertas presiones ambientales, se produce algún mutágeno que incrementa lo que se ha denominado el índice de mutación espontánea. Esto sería una clase de efecto semilamarckiano, aunque la *adaptación* seguiría procediendo sólo mediante la eliminación de las mutaciones; es decir, por selección natural. No habría, por supuesto, mucho que defender en esta conjetura, pues al parecer el índice de mutación espontánea basta para explicar la evolución adaptativa.¹²

He mencionado estos dos puntos sólo como advertencia contra la adhesión demasiado dogmática al darvinismo. Está claro que mi conjetura consiste en que el darvinismo está en lo correcto, hasta en el nivel del descubrimiento científico; y pienso que está en lo correcto aun más allá de este nivel: que está en lo justo incluso en el nivel de la creación artística. No descubrimos nuevos hechos o nuevos efectos copiándolos, ni infiriéndolos inductivamente, por observación; ni por ningún otro

¹² Entiendo que no se conocen mutágenos específicos (que actúen selectivamente, quizá en alguna particular secuencia de codones [un codón —derivado de código, latín, codex)— es, en el código genético, un triplete de bases que codifica cada uno de los aminoácidos que participan en la biosíntesis de proteínas. N. del T.], más que en otras secuencias de codones). Sin embargo, no sería asombrosa su existencia en este campo lleno de sorpresas; y podrían explicar "los puntos calientes" mutacionales. En todo caso, parece haber una verdadera dificultad en concluir, de la ausencia de mutágenos específicos conocidos, que no existen mutágenos específicos. Por tanto, me parece que el problema que se sugiere en el texto (la posibilidad de una reacción a ciertas presiones por la producción de mutágenos) está en pie.

método que proceda del medio biológico. Más bien utilizamos el método de prueba y eliminación del error. Como dice Ernst Gombrich, "el hacer es antes que el combinar":¹³ la producción activa de una nueva estructura tentativa ocurre antes de exponerla a las pruebas de eliminación.

vi

Yo sugiero, por tanto, que concibamos la manera de avanzar de la ciencia más o menos según los principios de las teorías de la formación de anticuerpos de Niels Jerne y sir Macfarlane Burnet.¹⁴ Las teorías anteriores relativas a la formación de anticuerpos suponían que el antígeno funcionaba como una plantilla negativa para la formación del anticuerpo. Esto significaría que habría una *enseñanza desde afuera*, desde el anticuerpo invasor. La idea fundamental de Jerne fue que la instrucción o información que permite al anticuerpo reconocer el antígeno es, literalmente, innata: que es parte de la estructura genética, aunque posiblemente esté sujeta a un repertorio de variaciones mutacionales. Transmite esta información el código genético, mediante los cromosomas de las células especializadas que producen los anticuerpos; y la reacción inmune es resultado de la estimulación de crecimiento que da a estas células el complejo anticuerpo-antígeno. Así, estas células son *seleccionadas* con la ayuda del medio invasor (es decir, con la ayuda del antígeno), y no mediante instrucciones. (La analogía con la selección —y con la modificación— de las teorías científicas está muy clara para Jerne, quien a este respecto se refiere a Kierkegaard y a Sócrates en el diálogo *Meno*.)

Con esta observación concluyo mi discusión de los aspectos biológicos del progreso en las ciencias.

¹³ Véase Ernst Gombrich, *Art and Illusion [Arte e ilusión]*, 1960, y ediciones posteriores, en el índice, en la entrada "making and matching".

¹⁴ Véase N. K. Jerne, "The Natural Selection Theory of Antibody Formation; Ten Years Later" ["La teoría de la selección natural en la formación de anticuerpos; diez años después"], pp. 301-312 de J. Cairns *et al.* (comps.), *Phage and the Origins of Molecular Biology [Phage y los orígenes de la biología molecular]*, 1966; también, "The Natural Selection Theory of Antibody Formation" ["La teoría de la selección natural en la formación de anticuerpos"], en *Proceedings of the National Academy of Sciences* 41, 1955, pp. 849-857; "Immunological Speculations" ["Especulaciones inmunológicas"], en *Annual Review of Microbiology* 14, 1960, pp. 341-358; "The Immune System" ["El sistema inmunitario"], en *Scientific American* 229, 1, 1973, pp. 52-60. Véase también: Macfarlane Burnet, "A Modification of Jerne's Theory of Anti-body Production, using the Concept of Clonal Selection" ["Una modificación de la teoría de Jerne de la producción de anticuerpos, utilizando el concepto de la selección clonal"], en *Australian Journal of Science* 20, 1957, pp. 67-69; y *The Clonal Selection Theory of Acquired Immunity [La teoría de la selección clonal de la inmunidad adquirida]*, 1959.

6. DOS CLASES DE DEFINICIONES (1945)

El peligro mayor para nuestra filosofía, aparte de la pereza y la falta de claridad, es el *escolasticismo*, [...] que está tratando lo que es vago como si fuera preciso...

F. P. RAMSEY

EL PROBLEMA de las definiciones y el del "significado de términos" constituyen la fuente más importante de la influencia intelectual de Aristóteles, que por desgracia aún prevalece; de todo ese escolasticismo verbal y vacío que se cierne como un fantasma, no sólo sobre la filosofía de la Edad Media, sino también sobre nuestra propia filosofía contemporánea; porque hasta una filosofía tan reciente como la de Wittgenstein padece, como veremos, de esta influencia. El desarrollo del pensamiento a partir de Aristóteles podría resumirse, en mi opinión, diciendo que cada disciplina, en tanto que utilizó el método aristotélico para definir, se ha quedado estancada en un estado de vacía verborrea y estéril escolasticismo, y que el grado en que las diversas ciencias han podido lograr algún progreso dependió del grado en que han sido capaces de apartarse de este método esencialista. (Ésta es la causa de que mucho de nuestra "ciencia social" pertenezca todavía a la Edad Media.) La discusión de este método tendrá que ser un tanto abstracta, debido a que Platón y Aristóteles embrollaron el problema a más no poder; y la influencia platónica y aristotélica ha suscitado tan hondos prejuicios, que la perspectiva de borrarlos no parece muy brillante. Pese a todo, acaso no carezca de interés analizar la fuente de tanta confusión y de tanta palabrería.

Aristóteles siguió a Platón en la distinción entre *conocimiento* y *opinión*.² Según Aristóteles, el conocimiento, o ciencia, puede ser de dos

¹ El epígrafe se ha tomado de *The Foundations of Mathematics [Los fundamentos de las matemáticas]*, 1931. [En relación con este capítulo, la atención del lector debe enfocarse especialmente a *Búsqueda sin término*, secciones 6 y 7.]

² En cuanto a la distinción que hace Platón, o más bien la que hace Parménides, entre conocimiento y opinión (distinción que siguió siendo muy del gusto de escritores más modernos, entre ellos, por ejemplo, Locke y Hobbes), véase *La sociedad abierta y sus enemigos*, capítulo 3, notas 22 y 26, y el texto mismo; la nota 19 al capítulo 5 y las notas 25-27 al capítulo 8. En cuanto a la correspondiente distinción que hace Aristóteles, véase, por ejemplo, *Metafísica*, 1039b31, y *Analítica posterior*, 88b30 y ss y 100b5.

clases: demostrativo o intuitivo. El *conocimiento demostrativo* también es un conocimiento de las "causas". Consiste en afirmaciones que pueden demostrarse —las conclusiones— junto con sus demostraciones silogísticas (que presentan las "causas" en sus "términos medios"). El *conocimiento intuitivo* consiste en captar "la forma indivisible" o esencia, o la naturaleza esencial de una cosa (si es "inmediata", es decir, si su "causa" es idéntica a su naturaleza esencial); es la fuente originaria de toda ciencia, ya que capta las premisas básicas originarias de todas las demostraciones.

Sin duda alguna, Aristóteles estaba en lo cierto cuando insistía en que no debemos intentar probar o demostrar *todo* nuestro conocimiento. Toda prueba debe proceder de premisas; la prueba, como tal, es decir, la derivación de las premisas, nunca puede, por tanto, establecer la verdad definitiva de ninguna conclusión, sino sólo mostrar que la conclusión debe ser verdadera *a condición* de que las premisas sean verdaderas. Si pudiéramos que las premisas debieran probarse a su vez, la cuestión de la verdad sólo retrocedería hacia otro paso o a un nuevo conjunto de premisas, y así sucesivamente, hasta el infinito. Fue para evitar tal regresión infinita (como dicen los lógicos) por lo que Aristóteles enseñaba que debemos suponer la existencia de premisas que son indudablemente verdaderas, y que no necesitan probarse en modo alguno; y a éstas las llamó "premisas básicas". Si damos por sentado que son buenos los métodos con los que derivamos conclusiones de estas premisas básicas, entonces podríamos decir, según Aristóteles, que todo el conocimiento científico está contenido en las premisas básicas, y que sería todo nuestro si pudiéramos obtener una lista enciclopédica de dichas premisas básicas. Pero, ¿cómo obtener estas premisas básicas? Como Platón, Aristóteles creía que en última instancia obtenemos todo el conocimiento mediante una intuitiva aprehensión de las esencias de las cosas. "Podemos conocer algo sólo si conocemos su esencia", dice

Para la distinción que hace Aristóteles entre conocimiento *demostrativo* y conocimiento *intuitivo*, véase el último capítulo de *Analítica posterior* (especialmente, 100b5-17; véase también 72b18-24, 75b31, 84a31, y 90a6-91a11). En cuanto a la conexión entre el conocimiento demostrativo y "las causas" de una cosa que son "distintas de su naturaleza esencial", y que por tanto necesitan un término medio, véase *op. cit.*, especialmente 93a5 y 93b26. En lo concerniente a la análoga conexión entre la intuición intelectual y la "forma indivisible" que capta —la esencia indivisible y la naturaleza individual que está identificada con su causa—, véase *op. cit.*, 72b24, 77a4, 85a1 y 88b35. Véase también *op. cit.*, 90a31: "Conocer la naturaleza de una cosa es conocer la razón por la que es" (es decir, su causa); y 93b21: "Son naturalezas esenciales las que son inmediatas, es decir, premisas básicas".

En cuanto al reconocimiento que hace Aristóteles de que debemos detenernos en alguna parte en la regresión de pruebas o demostraciones, y aceptar ciertos *principios* sin pruebas, véase, por ejemplo, *Metafísica*, 1006a7: "Es imposible probarlo todo, pues surgirá una regresión infinita..." Véase también *Analítica posterior*, 90b18-27.

Aristóteles, y agrega: "conocer una cosa es conocer su esencia". Según él, una "premisa básica" no es sino una afirmación que describe la esencia de una cosa. Pero tal afirmación es precisamente lo que él llama³ una definición. Por tanto, *todas las "premisas básicas" de pruebas son definiciones*.

¿A qué se parece una definición? Un ejemplo de definición sería: "Un perrito cachorro es un perro joven". El sujeto de tal frase definitoria, la expresión "perrito cachorro", es denominado *el término que hay que definir* (o el *término definido*); las palabras "perro joven" se denominan la *fórmula definidora*. Como regla general, la fórmula definidora es larga y más compleja que el término definido, y a veces mucho más. Aristóteles considera que el término que hay que definir⁴ es el nombre de la esencia de una cosa, y la fórmula definidora, la descripción de tal esencia. E insiste en que la fórmula definidora debe dar una descripción exhaustiva de la esencia o de las propiedades esenciales de la cosa que estamos definiendo; así, una afirmación como "Un cachorrito de perro tiene cuatro patas", aunque verdadera, no constituye una definición satisfactoria, puesto que no abarca todo lo que podemos llamar la

³ Véase *Metafísica*, 1031b7 y 1031b20. Véase también 996b20. "Una definición es una afirmación que describe la esencia de una cosa" (Aristóteles, *Tópicos*, 101b36, 153a, 153a15, etc. Véase también *Metafísica*, 1042a17); "La definición [...] revela la naturaleza esencial" (*Analítica posterior*, 91a1); "La definición es [...] una afirmación de la naturaleza de la cosa" (*ibid.*, 93b28); "Sólo tienen esencias aquellas cosas cuyas fórmulas son definiciones" (*Metafísica*, 1030a5 y ss); "La esencia, cuya fórmula es una definición, se llama también la sustancia de una cosa" (*ibid.*, 1017b21); "Entonces, resulta claro que la definición es la fórmula de la esencia..." (*ibid.*, 1031a13).

En lo que respecta a estos principios, es decir, a los puntos de partida o premisas básicas de las pruebas, debemos distinguir entre dos clases: 1) Los principios lógicos (véase *Metafísica*, 996b25 y ss), y 2) las premisas de las que deben proceder las pruebas, y que no pueden probarse a su vez si queremos evitar caer en una regresión infinita (véase la nota 2). Las últimas son definiciones: "Las premisas básicas o pruebas son definiciones" (*Analítica posterior*, 90b23; véase también 89a17 y 90a35). Asimismo, véase las pp. 45 y ss de W. D. Ross, *Aristotle*, 5ª edición, 1949, que comenta sobre la *Analítica posterior*, 73a20-74a4, "las premisas de la ciencia —escribe Ross (p. 46)— será, se nos dice, *per se* en cualquier sentido a) o b)". En la página anterior nos enteramos de que una premisa es necesaria *per se* (o esencialmente necesaria) en los sentidos a) y b) si se fundamenta en una definición.

⁴ "Si tiene un nombre, entonces habrá una fórmula de su significado", dice Aristóteles (*Metafísica*, 1030a14; véase también 1030b24); y explica que no toda fórmula del significado de un nombre es una definición; pero si el nombre es de una especie o de un género, entonces la fórmula será una definición.

Es importante anotar que, en mi utilización de la palabra "definición" (sigo aquí el uso moderno de esta palabra), se refiere siempre a toda la oración o a todo el enunciado de la definición, en tanto que Aristóteles (y otros que lo siguen en esto, por ejemplo, Hobbes) a veces también utiliza esta palabra como sinónimo de "*definiens*" [definidora, que define].

Las definiciones no son de particulares, sino sólo de universales (véase *Metafísica*, 1036a28) y sólo de esencias; es decir, de algo que es la especie de un género (es decir, una última *diferencia*; *ibid.*, 1038a19) y una forma indivisible; véase también *Analítica posterior*, 97b6 y s.

esencia de la "cachorrez", sino que esto también se aplica a un caballo; y de manera semejante, la declaración "Un cachorrito de perro es de color café", aunque puede ser verdad respecto de algunos, no es verdadera para todos los cachorritos de perro; y describe lo que es, no una esencia, sino meramente una propiedad accidental del término definido.

Pero la cuestión más difícil es cómo podemos aprehender definiciones o premisas básicas y estar seguros de que son las correctas, de que no hemos errado o captado la esencia errónea. Aunque Aristóteles no es muy claro acerca de este punto, no cabe duda de que, en lo principal, sigue el pensamiento de Platón. Platón pensaba⁵ que podemos captar las Ideas con el auxilio de una especie de infalible *intuición intelectual*; es decir, que las visualizamos o las vemos con nuestro "ojo mental", proceso que él concebía análogo a ver, pero dependiente sólo de nuestro intelecto, con exclusión de cualquier elemento que dependa de nuestros sentidos. El punto de vista de Aristóteles, a este respecto, es menos radical y menos inspirado que el de Platón, pero, a fin de cuentas, equivale a lo mismo.⁶ Porque, si bien enseña que llegamos a dar una definición sólo después de haber hecho muchas observaciones, reconoce que la experiencia sensorial por sí misma no capta la esencia universal, y que,

⁵ En cuanto a la doctrina de Platón al respecto, véase *La sociedad abierta y sus enemigos*, capítulo 8, sección iv.

Grote escribe, en la p. 260 de *Aristóteles*, 2ª edición, 1880: "Aristóteles había heredado de Platón esta doctrina de un infalible *Noûs* o Intelecto, que goza de completa inmunidad contra el error". Grote sigue haciendo hincapié en que, contrariamente al criterio de Platón, Aristóteles no desprecia la experiencia observacional, sino que asigna a su *Noûs* (es decir, a la intuición intelectual) "una posición como terminal y correlativa al proceso de la Inducción" (*loc. cit.*; véase también *op. cit.*, p. 577). Esto es así; pero la experiencia observacional sólo tiene, al parecer, la función de privilegiar y desarrollar nuestra intuición intelectual para su tarea, la intuición de la esencia universal; y, ciertamente, nadie ha explicado nunca cómo unas definiciones, *que están más allá del error*, pueden lograrse mediante la inducción.

⁶ El punto de vista de Aristóteles equivale al de Platón en cuanto a que, para ambos, en última instancia, no hay posibilidad de recurrir a la argumentación. Lo único que puede hacerse es aseverar *dogmáticamente* de una definición que es la verdadera descripción de su esencia; y se pregunta por qué esta descripción, y no otra, es la verdadera, lo único que nos queda es recurrir a "la intuición de la esencia".

Aristóteles habla de la inducción por lo menos en dos sentidos: en un sentido más heurístico, de un método que nos lleva a "intuir el principio general" (véase *Analítica anterior*, 67a22 y ss, 27b25-33, *Analítica posterior*, 71a7, 81a38-b5, 100b4 y ss), y en un sentido más empírico (*Analítica anterior*, 68b15-37, 69a16, *Analítica posterior*, 78a35, 81b5 y ss, *Tópicos*, 105a13, 156a4, 157a4).

En cuanto a "el cuerpo todo del hecho" mencionado en el siguiente párrafo, véase el final de la *Analítica posterior* (100b 15 y ss).

Es interesante observar cómo se parecen los puntos de vista de Hobbes (que era un nominalista, pero no un nominalista metodológico) al esencialismo metodológico de Aristóteles. Hobbes cree también que las definiciones son las premisas básicas de todo conocimiento (al contrario de la mera opinión).

por tanto, no puede determinar cabalmente una definición. Más adelante, simplemente postula que poseemos una intuición intelectual, una facultad mental o intelectual que nos permite captar infaliblemente la esencia de las cosas, y conocerlas. Y luego considera que, si sabemos algo, si conocemos una esencia intuitivamente, debemos ser capaces de describirla y, por tanto, de definirla. (Sus argumentos en la *Analítica posterior* en favor de esta teoría son asombrosamente débiles. Consisten sólo en señalar que nuestro conocimiento de las premisas básicas no puede ser demostrativo, pues esto nos llevaría a una regresión infinita, y que las premisas básicas deben ser por lo menos tan verdaderas y ciertas como las conclusiones que se basan en ellas. Escribe Platón: "De esto se colige que no puede haber conocimiento demostrativo de las premisas primeras; y puesto que nada, sino la intuición intelectual, puede ser más verdadero que el conocimiento demostrativo, de esto se colige que debe ser la intuición intelectual la que capta o aprehende las premisas básicas". En *De Anima* y en la parte teológica de la *Metafísica*, encontramos más de un argumento de este tenor; porque estamos aquí frente a una *teoría* de la intuición intelectual, que entra en contacto con su objeto, la esencia, y que incluso se confunde o se auna con su objeto: "El conocimiento real es idéntico a su objeto".)

Para resumir este breve análisis, creo que podemos ofrecer una buena descripción del ideal aristotélico del conocimiento completo y perfecto si decimos que Aristóteles consideró la meta última de toda investigación compilar una enciclopedia que contuviera las definiciones intuitivas de todas las esencias, es decir, sus nombres, junto con sus fórmulas definidoras o definitorias; y que él consideraba que el progreso del conocimiento consistía en la gradual acumulación de tal enciclopedia, en ampliarla y en llenar las lagunas que hubiera en ella, y por supuesto, en la derivación silogística, a partir de ella, "de todo el cuerpo de hechos" que constituye el conocimiento demostrativo.

Ahora bien, no cabe la menor duda de que todos estos puntos de vista esencialistas se ponen en el máximo contraste posible respecto a los métodos de la ciencia moderna. (Pienso, al afirmar esto, en las ciencias empíricas; quizá no en las matemáticas puras.) Porque, ante todo, aunque en materia de ciencia nos esforzamos al máximo en descubrir la verdad, estamos conscientes de que nunca estaremos seguros de haberla descubierto. Hemos aprendido en el pasado, a costa de muchas decepciones, que no debemos esperar ninguna finalidad. Y hemos aprendido a ya no sentirnos desilusionados si son derrumbadas nuestras teorías científicas; porque podemos, en la mayoría de los casos, determinar con gran confianza cuál, entre dos teorías, es la mejor. Por consiguiente, podemos saber que estamos realizando progresos; y es este conocimiento

el que, para la mayoría de nosotros, nos compensa de la pérdida de las ilusiones de finalidad y certidumbre. En otras palabras, sabemos que nuestras teorías científicas deben seguir siendo, siempre, hipótesis. Pero sabemos también que, en muchos casos importantes, podemos averiguar si una nueva hipótesis es superior a otra vieja hipótesis. Porque, si difieren, entonces nos llevarán a hacer diferentes predicciones, que a menudo pueden ponerse a prueba experimentalmente; y, basados en ese experimento crucial, podemos darnos cuenta de que la nueva teoría nos lleva a resultados satisfactorios, allí donde la vieja fracasa. Así, podemos decir que, en nuestra búsqueda de la verdad, hemos sustituido la certidumbre científica con el progreso científico. Y este enfoque del método científico está corroborado por el desarrollo de la ciencia. Porque la ciencia no se desarrolla mediante una gradual acumulación enciclopédica de información esencial, como pensaba Aristóteles, sino mediante un método mucho más revolucionario; la ciencia avanza mediante ideas audaces, mediante el avance de nuevas y extravagantes teorías (como la de que la Tierra no es plana, o la de que el "espacio métrico" no es plano), y mediante el derrumbe de las viejas ideas y teorías.

Pero este enfoque del método científico [que se desarrolla más adelante, en los escritos 9-14] significa que en la ciencia no existe el "conocimiento", en el sentido en que Platón y Aristóteles entendían esta palabra, en el sentido de finalidad; en ciencia, jamás tenemos razón suficiente para pensar que hemos alcanzado la verdad. Lo que solemos llamar "conocimiento científico" es, generalmente, no "conocimiento" en este sentido, sino más bien información sobre las diversas hipótesis que compiten y sobre la manera como han soportado varias pruebas; es, para emplear el lenguaje de Platón y de Aristóteles, información acerca de la última y mejor probada "opinión" científica. Además, este punto de vista significa que en ciencia no tenemos pruebas (excepto, por supuesto, en matemáticas puras y en lógica). En las ciencias empíricas, que son las únicas capaces de darnos información sobre el mundo en que vivimos, las pruebas no se producen, si atribuimos a la palabra "prueba" el significado de un argumento que establece de una vez por todas, y para siempre, la verdad de alguna teoría. (Lo que sí puede ocurrir, sin embargo, son las refutaciones de las teorías científicas.) Por otra parte, las matemáticas puras y la lógica, que sí permiten las pruebas, no nos dan información acerca del mundo, sino que sólo desarrollan los medios que lo describen. Así pues, podemos afirmar (como ya lo he señalado en otro lugar):⁷ "En la medida en que una afir-

⁷ La cita es de mi nota en *Erkenntnis* 3, 1933, al pie de la p. 426, ahora en su traducción, *The Logic of Scientific Discovery* [La lógica de la investigación científica], pp. 312-314;

mación científica habla acerca de la realidad, debe ser falsificable; y en la medida en que no sea falsificable, no habla acerca de la realidad." Pero aunque la prueba no tiene parte alguna en las ciencias empíricas, los argumentos sí la tienen; en realidad, desempeñan un papel por lo menos tan importante como el jugado por la observación y la experimentación.

El papel de las definiciones, en ciencia sobre todo, también es muy diferente de lo que pensaba Aristóteles al respecto. Aristóteles enseñaba que en una definición hemos señalado primero la esencia —quizás al nombrarla— y que luego la describimos con la ayuda de la fórmula definitoria; así como en una frase ordinaria como "Este cachorrito de perro es café", primero señalamos cierta cosa diciendo "este cachorrito de perro", y luego la describimos como "de color café". Y enseñaba que, al describir así la esencia hacia la que señala el término que hay que definir, así también determinamos o explicamos el *significado* del término. Según esto, la definición puede al mismo tiempo contestar a dos preguntas muy estrechamente relacionadas. Una es: "¿Qué es eso?" (por ejemplo: "¿Qué es un cachorrito de perro?"); la otra es "¿Qué significa?" (por ejemplo, "¿Qué significa un cachorrito de perro?"); pregunta por el significado de un término (a saber, el término que denota la esencia). En el presente contexto, no es necesario distinguir entre estas dos preguntas; es más importante ver qué tienen en común; y deseo, especialmente, llamar la atención sobre el hecho de que *ambas preguntas son suscitadas por el término que está, en la definición, en el lado izquierdo, y son contestadas por la fórmula definitoria que está en el lado derecho*. Este hecho caracteriza al enfoque esencialista, del que difiere radicalmente el método científico para hacer definiciones.

Aunque podemos decir que la interpretación esencialista lee una definición "normalmente", es decir, *de izquierda a derecha*, podemos afirmar también que una *definición*, tal como se *utiliza normalmente en la ciencia moderna, debe leerse de abajo hacia arriba, o de derecha a izquierda*; porque empieza con la fórmula definitoria, y pide una breve etiqueta para ella. Así, el punto de vista científico de la definición "Un cachorrito de perro es un perro joven" sería que es una respuesta a la pregunta "¿Cómo llamaremos a un perro joven?", más que a la pregunta "¿Qué es un cachorrito de perro?" (Las preguntas como "¿Qué es la vida?" o "¿Qué es la gravedad?" no desempeñan ningún papel en la ciencia.) La utilización científica de las definiciones, caracterizada por el enfoque "de derecha a izquierda", puede llamarse su interpretación *nominalista*,

es una variante y generalización de la aseveración acerca de la geometría que hizo Einstein en "Geometry and Experience", 1921, pp. 232-246 de *Ideas and Opinions*.

⁸ Véase, por ejemplo, *Metafísica*, 1030a6 y 14 (véase la nota 4, más arriba).

opuesta a su interpretación aristotélica o *esencialista*.⁹ En la ciencia moderna, sólo¹⁰ se dan las definiciones nominalistas; es decir, que se introducen símbolos abreviados o de taquigrafía, etiquetas, para acortar algo que sería de prolija exposición en lenguaje normal. Y en seguida podemos advertir que las definiciones *no* desempeñan ningún papel muy importante en la ciencia. Porque, por supuesto, los símbolos abreviados pueden sustituirse siempre por las expresiones largas, las fórmulas definitorias, a las que representan. En algunos casos, hacerlo así convertiría nuestro lenguaje científico en algo muy incómodo, que ocasionaría gran desperdicio de tiempo y de papel. Pero nunca deberíamos perder el menor fragmento de información objetiva. Nuestro "conocimiento científico", en el sentido en que puede emplearse apropiadamente esta expresión, seguirá enteramente igual si prescindimos de todas las definiciones;

⁹ Deseo recalcar que hablo aquí de *nominalismo contra esencialismo* de manera puramente metodológica. No tomo ninguna posición respecto del problema *metafísico* de los universales, es decir, respecto del problema metafísico del nominalismo contra el esencialismo (término que, sugiero, debería emplearse en vez del tradicional término "realismo"); y ciertamente no defiendo el nominalismo metafísico, aunque sí estoy en favor del nominalismo metodológico. [Véase también la nota 13 al escrito 4, más arriba.]

La oposición entre las *definiciones nominalistas y esencialistas* que se plantea en el texto es un intento de reconstruir la distinción tradicional entre las definiciones "verbales" y las "reales". *Mi principal insistencia, sin embargo, es en la cuestión de si la definición se lee de derecha a izquierda o de izquierda a derecha; o, en otras palabras, si sustituye una larga "historia" con una breve, o una breve con una larga.*

¹⁰ Mi advertencia de que en ciencia *sólo* se dan las definiciones nominalistas (hablo aquí de definiciones explícitas, y no de definiciones implícitas ni de definiciones recursivas) necesita alguna defensa. No implica, por cierto, que en ciencia no se utilicen los términos más o menos "intuitivamente"; esto está claro si sólo consideramos que todas las cadenas de definiciones deben empezar con términos *no definidos*, cuyo significado puede ejemplificarse, pero no definirse. Además, en ciencia, especialmente en matemáticas, resulta claro que a menudo utilizamos un término, por ejemplo, "dimensión" o "verdad", intuitivamente, pero luego procedemos a definirlo. Pero esto constituye una descripción más bien burda de la situación. Una descripción más precisa sería la siguiente: Algunos de los términos no definidos utilizados intuitivamente pueden sustituirse, a veces, con términos definidos, de los que puede demostrarse que satisfacen las intenciones con las que se han utilizado los términos no definidos; es decir, en cada enunciado en que aparecieron los términos no definidos (por ejemplo, lo que se interpretó como analítico) hay un enunciado correspondiente en que aparece el término nuevamente definido, o redefinido (el cual se colige de la definición).

Bien podríamos decir que K. Menger ha definido recursivamente la palabra "dimensión", o que A. Tarski ha definido el término "verdad"; pero esta manera de expresarnos puede suscitar malas interpretaciones; lo que ha sucedido es que Menger dio una definición puramente nominal de clases de conjuntos de puntos que él llamó "n-dimensionales", porque era posible sustituir el concepto matemático intuitivo "n-dimensional" con un nuevo concepto en contextos de gran importancia; y lo mismo podemos afirmar del concepto de Tarski de "la verdad". Tarski dio una definición nominal (o, más bien, un método para hacer definiciones nominales) a la que etiquetó como "verdad", ya que un sistema de enunciados podría derivarse de la definición correspondiente a esos enunciados (como la ley del término medio excluido), y que habían utilizado muchos lógicos y filósofos en relación con lo que ellos habían llamado "verdad".

el único efecto que tienen es que afectan a nuestro lenguaje, que con ello no perdería precisión, sino sólo brevedad. (Esto no debe interpretarse como si en la ciencia no hubiese la urgente necesidad práctica de introducir definiciones, en obsequio a la brevedad.) No podría haber mayor contraste entre este enfoque del papel de las definiciones y del punto de vista de Aristóteles. Porque las definiciones esencialistas de Aristóteles son los principios de los que deriva todo nuestro conocimiento; por tanto, contienen todo nuestro conocimiento, y sirven para sustituir una fórmula muy larga con una breve. Opuestas a este enfoque, las definiciones científicas o nominalistas no contienen conocimiento alguno, y ni siquiera alguna "opinión"; lo único que hacen es introducir nuevas y arbitrarias etiquetas abreviadas; convierten una exposición larga en breve síntesis.

En la práctica, estas etiquetas son muy útiles. Para darnos cuenta de ello, basta que consideremos las gravísimas dificultades que se suscitarían si un bacteriólogo, siempre que hablara de cierta cepa de bacterias, tuviera que repetir cada vez toda su descripción (incluyendo los métodos de tinción, etc., mediante la cual se distingue de otras varias especies similares). Y, con una consideración semejante a la anterior, podemos entender por qué se ha olvidado tan a menudo, incluso por los científicos, que las definiciones científicas deben leerse "de derecha a izquierda", como ya lo hemos explicado. Porque la mayoría, cuando se inicia en el estudio de una ciencia (por ejemplo, la bacteriología), debe tratar de averiguar los significados de todos esos términos técnicos nuevos a los que se enfrenta. De esta manera, el estudiante en realidad *aprende* la definición "de izquierda a derecha", sustituyendo, como si se tratara de una definición esencialista, una exposición muy larga con otra muy breve. Pero esto es sólo un accidente psicológico, y el maestro, o el autor de un libro de texto, debe proceder de manera muy diferente: es decir, que introducirá un término técnico sólo cuando surja la necesidad de hacerlo.

Hasta aquí he intentado demostrar que el empleo científico o nominalista de las definiciones es del todo diferente del método esencialista aristotélico. Pero también podemos demostrar que el punto de vista esencialista en cuanto a las definiciones es sencillamente insostenible en sí mismo. Para no prolongar demasiado esta discusión, criticaré únicamente dos de las doctrinas esencialistas; dos doctrinas que revisten alguna importancia, porque algunas escuelas modernas influyentes todavía se basan en ellas. Una es la doctrina esotérica de la intuición intelectual; la otra, la muy popular doctrina de que "debemos definir nuestros términos", si deseamos ser precisos.

Aristóteles sostenía, siguiendo a Platón, que poseemos una facultad, la intuición intelectual, mediante la cual podemos visualizar esencias y des-

cubrir cuál definición es la correcta; y muchos esencialistas modernos han repetido esta doctrina. Otros filósofos, siguiendo a Kant, sostienen que no poseemos tal facultad. Mi opinión es que podemos aceptar sin objeción que poseemos algo que podríamos describir como "intuición intelectual"; o, más precisamente, que algunas de nuestras experiencias intelectuales pueden describirse así. Todo aquel que "entiende" una idea, o un punto de vista, o un método aritmético (por ejemplo, la multiplicación), en el sentido de que "ha captado o sentido cómo se hace", podría decir que entiende eso intuitivamente; y existen incontables experiencias o vivencias intelectuales de esa clase. Pero yo insistiría, por otra parte, en que estas experiencias, por importantes que sean para nuestras actividades científicas, nunca deben servir para establecer la verdad de una idea o teoría, por muy fuerte que sea la sensación intuitiva de alguien que así se convenza de que debe ser verdadera, o de que es "evidente por sí misma".¹¹ Tales intuiciones tampoco pueden servir de argumento, aunque nos animen a buscar argumentos. Porque alguien más podría tener una intuición igualmente fuerte en el sentido de que esa misma teoría es falsa. El camino de la ciencia está pavimentado con teorías descartadas que alguna vez se consideraron evidentes por sí mismas. Francis Bacon, por ejemplo, se mofaba de quienes negaban la evidente verdad de que el Sol y las estrellas giraban alrededor de la Tierra, que obviamente estaba fija. Sin duda, la intuición desempeña un gran papel en la vida de un científico, como sucede en la vida de un poeta. Lo

¹¹ El hecho de que una aseveración sea verdadera puede ayudarnos, a veces, a explicar por qué nos parece evidente por sí misma. Tal es el caso de " $2 + 2 = 4$ ", o del enunciado "El sol irradia luz y calor". Pero lo contrario claramente no se ajusta a este caso. El hecho de que una frase o un enunciado nos parezca a algunos, o a todos, evidente por sí mismo, es decir, que algunos o todos creamos firmemente en su veracidad, y no podamos concebir su falsedad, no es razón suficiente para que sea veraz. (El que no podamos concebir la falsedad de una aseveración es en muchos casos sólo una razón para sospechar que nuestro poder de imaginación es deficiente o que no se ha desarrollado.) Constituiría uno de los más graves errores que una filosofía nos presentara la evidencia en sí misma como en favor de la verdad de una frase o de un enunciado; sin embargo, es esto lo que hacen casi todas las filosofías idealistas. Nos demuestra que las filosofías idealistas son a menudo sistemas de apología de algunas creencias dogmáticas.

La excusa de que a menudo estamos en una posición tal que debemos aceptar ciertos enunciados o ciertas frases por la sola razón de que son evidentes por sí mismos, no es válida. Los principios de la lógica y del método científico (especialmente el principio de "inducción", o el de "la ley de la uniformidad de la naturaleza") se expresan generalmente como aseveraciones que debemos aceptar, y que no podemos justificar sino con su evidencia en sí misma. Aun si esto fuera cierto, sería más franco de nuestra parte decir que no podemos justificar tales aseveraciones sin más. Pero, de hecho, no hay necesidad de un "principio de inducción". [Véase el siguiente escrito.] Y en lo que se refiere a los "principios de la lógica", mucho se ha hecho en años recientes para demostrar que la teoría de la evidencia por sí o en sí misma es obsoleta. (Véase, especialmente, R. Carnap, *Logical Syntax of Language* [Sintaxis lógica del lenguaje], 1937, y su *Introduction to Semantics*, 1942.)

conduce hacia sus descubrimientos; pero también puede conducirlo hacia sus fracasos. Y siempre seguirá siendo esta intuición su asunto privado, por así decirlo. La ciencia no pregunta cómo obtuvo el científico sus ideas, pues sólo le interesan los argumentos que todos podemos poner a prueba. El gran matemático Cari Friedrich Gauss describió muy claramente esta situación cuando exclamó en una ocasión: "¡Ya obtuve mi resultado; pero todavía no sé cómo obtenerlo!" Todo esto se aplica, por supuesto, a la doctrina aristotélica de la intuición intelectual de las llamadas esencias, que propagaron Hegel y, en nuestro propio tiempo, E. Husserl y sus numerosos discípulos; e indica que "la intuición intelectual de las esencias" o "fenomenología pura", como la llama Husserl, es un método que no pertenece ni a las ciencias ni a la filosofía. (La muy debatida cuestión de si se trata de una nueva invención, como piensan los partidarios de la fenomenología pura, o quizás una versión del cartesianismo o del hegelianismo, puede dirimirse fácilmente: es una versión del aristotelismo.)

La segunda doctrina que voy a criticar tiene incluso más importantes conexiones con los enfoques modernos de la ciencia; y tiene que ver especialmente con el problema del verbalismo. Desde Aristóteles, se ha sabido y aceptado generalmente que no podemos probar todas las afirmaciones, y que el intento de hacerlo sería vano, pues sólo nos llevaría a una regresión infinita de las pruebas. Pero ni Aristóteles¹² ni, al parecer, muchos escritores modernos parecen darse cuenta de que el intento análogo de definir el significado de todos nuestros términos debe llevarnos, de igual manera, a una infinita regresión de definiciones. El siguiente pasaje de *Platón hoy*, obra de Crossman, es característico de un punto de vista que, por implicación, sostienen muchos filósofos contemporáneos famosos, entre ellos, por ejemplo, Wittgenstein:¹³

¹² "La ciencia supone las definiciones de todos sus términos..." (Ross, *Aristóteles*, p. 44; véase *Analítica posterior*, 76a32-36); véase también, más arriba, la nota 4.

¹³ R. H. S. Crossman, *Plato To-day [Platón, hoy]*, 1937, 2ª edición, 1959, p. 51.

Una doctrina muy parecida a ésta la expresan M. R. Cohn y E. Nagel en su libro *An Introduction to Logic and Scientific Method [Introducción a la lógica y al método científico]*, 1934, p. 232: "Muchas de las disputas acerca de la verdadera naturaleza de la propiedad, de la religión, de la ley [...] seguramente desaparecerían si los equivalentes definidos con precisión sustituyeran a estas palabras". (Véase también la nota 15, más adelante.)

Los puntos de vista respecto a este problema que expresó Wittgenstein en su *Tractatus*, 1921/1922, y que también expresaron varios seguidores suyos, no son tan definidos como los de Crossman, Cohn y Nagel. Wittgenstein es un antimetafísico. Escribe en el Prefacio: "Este libro trata de los problemas de la filosofía y demuestra, en mi opinión, que el método para formular estos problemas depende de la falta de comprensión de la lógica de nuestro lenguaje". Trata de demostrar que la metafísica es "simplemente absurda", e intenta trazar un límite, en nuestro lenguaje, entre lo sensato y lo insensato, entre lo que tiene sentido y lo que no lo tiene: "el límite puede [...] trazarse, en los lenguajes, y lo que ha

...si no sabemos con toda precisión el significado de las palabras que utilizamos, no podemos discutir de nada en forma provechosa. La mayoría de los argumentos fútiles en los que todos perdemos el tiempo se deben en gran medida a que cada uno de nosotros tiene su propio significado de las palabras que utiliza, y damos por sentado que nuestros oponentes las utilizan con el mismo significado. Si empezáramos por definir claramente nuestros términos, podríamos tener discusiones mucho más provechosas. Como un ejemplo de esto, con sólo leer los diarios podemos observar que la propaganda política (la contraparte moderna de la retórica) depende en gran medida, para tener éxito, de confundir el significado de los términos. Si la ley obligara a los políticos a definir todos los términos que desearan utilizar, perderían gran parte de su atractivo popular, sus discursos serían más breves, y se vería que muchos de sus desacuerdos son puramente verbales.

Este pasaje es muy característico de uno de los prejuicios que debemos a Aristóteles, del prejuicio de que el lenguaje puede hacerse más preciso mediante el empleo de las definiciones. Veamos si realmente puede hacerse esto.

Ante todo, podemos ver claramente que si a "los políticos" (o a alguien más) "los obligara la ley a definir todos los términos que desearan utilizar", sus discursos no serían más breves, sino infinitamente largos. Porque una definición no puede establecer el significado de un término, así como una derivación lógica no puede establecer la verdad de una afirmación; ambas sólo pueden causar la regresión del problema. La derivación traslada el problema de la verdad de regreso a sus premisas, y la definición traslada el problema del significado de regreso a los términos definitorios (es decir, a los términos que constituyen la fórmula definitoria).¹⁴ Pero estos términos, por muchas razones, seguramente

quedado en el otro lado del límite será simplemente tontería, absurdidad". Según este libro de Wittgenstein, las proposiciones sí tienen algún sentido válido. Son verdaderas o falsas. Las proposiciones filosóficas no existen; sólo parecen proposiciones, pero en realidad son absurdidades. El límite entre lo que tiene significado o sentido y lo que no lo tiene coincide con el límite entre las ciencias naturales y la filosofía: "La totalidad de las verdaderas proposiciones es la ciencia natural total (o la totalidad de las ciencias naturales). La filosofía no es la filosofía de las ciencias naturales". *Por tanto, la verdadera tarea de la filosofía no es formular proposiciones; más bien consiste en esclarecer proposiciones*: "El resultado de la filosofía no es un número de *proposiciones filosóficas*, sino hacer que sean claras las proposiciones". Quienes no vean esto, y formulen proposiciones filosóficas, hablan de tonterías metafísicas.

¹⁴ Es importante distinguir entre una deducción lógica, en general, y una prueba o demostración, en particular. Una *prueba* o *demostración* es un argumento deductivo con el que se establece finalmente la verdad de la conclusión; así es como utiliza Aristóteles este término, exigiendo (por ejemplo, en *Analítica posterior*, 73a20 y ss) que la verdad "necesaria" de la conclusión se establezca; y así es como Carnap utiliza el mismo término (véase, especialmente, la *Lógica! Syntax of Language*, sección 10, p. 29, sección 47, p. 171), en que demuestra que las conclusiones que son "demostrables" en este sentido son "analíticamente" verdaderas. (No me ocuparé aquí de los problemas relativos a los términos "analítico" y "sintético".)

serán tan vagos y confusos como los términos con los que empezamos nuestra demostración; y, en todo caso, debemos proseguir hasta definirlos, a su vez. Y así sucesivamente, hasta el infinito. Vemos así que la exigencia de que se definan bien todos los términos es tan insostenible como la de que se comprueben todas nuestras afirmaciones.

A primera vista, esta crítica puede parecer injusta. Se nos puede objetar que lo que la gente desea, si exige definiciones, es la eliminación de las ambigüedades tan a menudo relacionadas con palabras como¹⁵ "democracia", "libertad", "deber", "religión", etc.; que está claro que resulta imposible definir todos nuestros términos, pero que sí es posible definir algunos de los más peligrosos, y limitarnos a eso; y que los términos definitorios deben aceptarse; es decir, que debemos detenernos después de uno o dos pasos en este sentido, para evitar caer en una regresión infinita. Esta defensa, sin embargo, es insostenible. Es verdad que los términos mencionados se utilizan muy mal. Pero niego que el intento de definirlos mejore la situación; más bien sólo la empeoraría. Está claro que "definiendo sus términos", aunque fuera una sola vez, y dejando sin definir sus términos definitorios, los políticos no podrían acortar sus discursos; porque cualquier definición esencialista, es decir, una que "defina nuestros términos" (opuesta a una nominalista, que introduce nuevos tecnicismos), significa sustituir una exposición larga con una breve, como hemos visto. Además, el intento de definir términos sólo incrementaría la vaguedad y la confusión. Porque si no podemos exigir que todos los términos definitorios se definan a su vez, un político o filósofo astuto podría fácilmente satisfacer la exigencia de definiciones. Si se le preguntara, por ejemplo, qué entiende por "democracia", podría contestar: "el gobierno de la voluntad general", o "el gobierno del

espíritu del pueblo"; y como así ha dado una definición y ha cumplido con las más altas normas de la precisión, ya nadie se atreverá a criticarlo en lo sucesivo. Y, ciertamente, ¿cómo se le podría criticar, ya que la exigencia de que definiera a su vez "gobierno", o "pueblo", o "voluntad", o "espíritu" nos pondría en camino de una regresión infinita, que todos vacilarían en suscitarse? Pero, si pese a todo se suscitara tal regresión, fácilmente se podría satisfacer la exigencia de definición. Por otra parte, la controversia acerca de si la definición dada era correcta, o ver-

Podemos añadir aquí algunas observaciones generales acerca de la inutilidad de las definiciones esencialistas:

1) El intento de resolver un problema fáctico refiriéndonos a definiciones suele significar la sustitución de un problema meramente verbal en vez del problema de hechos. (Hay un excelente ejemplo de este método en la *Física* de Aristóteles, 197b6-32.) Esto se puede demostrar con los siguientes ejemplos: a) Hay un problema fáctico: ¿Podemos retornar a la jaula del tribalismo? ¿Con qué medios? b) Hay un problema moral: ¿Deberíamos retornar a esa jaula?

El filósofo del significado, si se le enfrenta a a) o a b), dirá: "Todo depende de lo que quiera usted decir con sus términos vagos; dígame cómo define usted *retornar*, *jaula* y *tribalismo*, y con ayuda de estas definiciones, *podría yo ser capaz de resolver su problema*". Sostengo, en contra de esto, que si la decisión o resolución del problema puede hacerse con el auxilio de las definiciones, si se deriva de las definiciones, entonces el problema que así se dirime era sólo un problema verbal; pues se ha resuelto independientemente de los hechos o de las decisiones morales.

2) El filósofo *esencialista* del significado quizá diga algo peor, especialmente en lo que respecta al problema b); podría sugerir, por ejemplo, que depende de "la esencia", o del "carácter esencial", o quizá del "destino" de nuestra civilización si debiéramos o no intentar ese retorno.

3) El esencialismo y la teoría de la definición nos han conducido hacia un asombroso desarrollo en ética. Este desarrollo consiste en la creciente abstracción y la pérdida de contacto con la *base de toda ética*; los problemas morales prácticos, que tenemos que resolver aquí y ahora. Nos lleva a hacernos la pregunta general: "¿Qué es bueno?", o "¿Qué es lo bueno?"; luego, tendremos que preguntar: "¿Qué significa *bueno*?", y por último preguntaremos: "¿Puede resolverse el problema que plantea la pregunta '¿Qué significa *bueno*?'", o bien: "¿Puede definirse la palabra *bueno*?" G. E. Moore, que planteó este último problema en su libro *Principia Ethica*, 1903, estaba en lo cierto cuando insistía en que "bueno", en la acepción moral, no puede definirse en términos "naturalistas". Porque, si pudiéramos definirlo así, significaría algo como "amargo", o "dulce", o "verde", o "rojo"; y entonces estaría muy fuera de lugar desde el punto de vista de la moralidad. Así como no tenemos que lograr lo amargo, o lo dulce, etc., no habría ninguna razón para que tuviéramos algún interés moral en "lo bueno". Pero, aunque Moore estaba en lo cierto en lo que acaso se considera justamente su tesis principal, podemos sostener que el análisis de lo bueno o de cualquier otro concepto o esencia de ninguna manera puede contribuir a elaborar una teoría ética que se refiera a la única base pertinente de toda la ética, el problema moral inmediato que debe resolverse aquí y ahora. Tal análisis sólo nos puede llevar a sustituir un problema moral con otro meramente verbal. (Véase también la nota 18(1) al capítulo 5 de *La sociedad abierta y sus enemigos*, especialmente acerca de lo improcedente de los juicios morales.) [En esta nota, Popper critica la ética "científica" tachándola de absoluta esterilidad. Y añade: "...Pero es evidente que los *juicios morales* no vienen al caso (son improcedentes). Sólo a los chismosos les interesa juzgar a la gente [...] La importancia suprema de los problemas morales estriba, por supuesto, en que podemos actuar con previsión inteligente y en que siempre podemos preguntarnos cuáles deben ser nuestros objetivos..." (N. del T.)]

Desde los tiempos de Aristóteles se ha establecido claramente que no todas las deducciones lógicas son pruebas (es decir demostraciones); también existen deducciones lógicas que no son pruebas; por ejemplo, podemos deducir conclusiones a partir de premisas reconocidas como falsas, y a estas deducciones no las llamamos "pruebas". Carnap llama a las deducciones no demostrativas "derivaciones" (*loe. cit.*). Resulta interesante que no se haya asignado antes un nombre a estas deducciones no demostrativas; demuestra la preocupación por las pruebas, que surgió del prejuicio aristotélico de que la "ciencia" o el "conocimiento científico" debe establecer bien todas sus premisas, es decir, aceptarlas como evidentes por sí mismas, o probarlas. Pero la posición es ésta: *Fuera de la lógica pura y de las matemáticas puras, nada puede probarse*. Los argumentos en otras ciencias (e incluso algunos en matemáticas, como ha demostrado I. Lakatos) no son pruebas, sino sólo *derivaciones*.

Podemos observar que existe un paralelismo de largo alcance entre los problemas de la *derivación*, por un lado, y la *definición*, por el otro, y entre problemas de la *veracidad de los enunciados*, y el del *significado de los términos*. [Véase, en especial, más arriba, el cuadro de la p. 79.]

¹⁵ Los ejemplos son los mismos que recomiendan Cohén y Nagel, *op. cit.*, pp. 232 y ss, en cuanto a la definición. (Véase la nota 13, más arriba.)

dadera, sólo nos llevaría a otra controversia hueca sobre el significado de las palabras.

Así, cae por tierra el punto de vista esencialista de la definición, incluso si no intenta, siguiendo a Aristóteles, establecer los "principios" de nuestro conocimiento, sino sólo plantea la exigencia, aparentemente más modesta, de que "deberíamos definir el significado de nuestros términos".

Mas no cabe la menor duda de que la exigencia de que hablemos claramente y sin ambigüedades resulta importantísima, y debemos cumplirla. ¿Puede el punto de vista nominalista satisfacerla? Y, ¿puede el nominalista escapar de la regresión infinita?

Sí, puede hacerlo. Porque para la posición nominalista no hay dificultad que corresponda a caer en una regresión infinita. Como hemos visto, la ciencia no utiliza definiciones para determinar el significado de sus términos, sino sólo para introducir etiquetas de taquigrafía práctica. Y no depende de las definiciones; todas las definiciones pueden omitirse sin perderse la información que se da. De esto se desprende que, en la ciencia, *todos los términos realmente necesarios deben ser términos no definidos*. Entonces, ¿cómo aseguran las ciencias el significado de sus términos? Se han sugerido varias respuestas para esta pregunta,¹⁶ pero no creo que alguna de ellas sea satisfactoria. La situación parece ser ésta: el aristotelismo y las filosofías relacionadas con él nos han dicho desde hace tanto tiempo que es tan importante obtener un conocimiento preciso del significado de nuestros términos, que todos nos inclinamos por creer esto. Y seguimos aferrados a esta creencia, a pesar del incuestionable hecho de que la filosofía, que durante 20 siglos se ha preocupado por el significado de sus términos, no sólo está llena de verbalismo, sino que sigue siendo abrumadoramente vaga y ambigua, mientras que una ciencia como la física, que apenas se preocupa por los términos y su significado, y sólo se concentra en los hechos, ha logrado llegar a una gran precisión. Esto, con toda seguridad, debe tomarse como indicio de que, bajo la influencia de Aristóteles, la importancia del significado de los términos se había exagerado mucho. Pero creo

¹⁶ Pienso al respecto en los métodos de "constitución", "definición implícita", "definición por correlación" y "definición operacional". Los argumentos de los "operacionalistas" parecen, en general, bastante sólidos; pero no pueden pasar por alto el hecho de que en sus definiciones operacionales, o descripciones, necesitan términos universales que han de aceptarse como indefinidos; y el problema sigue aplicándoseles. A pesar de esto, podemos operar con esos términos cuyo significado hemos aprendido "operacionalmente". Los utilizamos, por así decirlo, de manera que nada dependa, o dependa muy poco, de su significado, cuanto sea esto posible. Nuestras "definiciones operacionales" tienen la ventaja de ayudarnos a desviar el problema a un campo en el que nada depende, o depende muy poco, de las palabras. *Hablar claro es hablar de tal manera que las palabras mismas no tengan importancia.*

que indica algo más: porque no sólo esta concentración en el problema del significado fracasa en su intento de establecer la precisión; es en sí misma una fuente de vaguedad, ambigüedad y confusión.

En la ciencia, tenemos cuidado de que las afirmaciones que hacemos nunca *dependan* del significado de nuestros términos. Incluso cuando los términos están bien definidos, jamás tratamos de hacer derivar ninguna información de la definición misma, ni de fundamentar un argumento en esa definición. Por esta razón, los términos que utilizamos no nos plantean problema alguno. No los sobrecargamos. Tratamos de que tengan el menor peso posible. No tomamos demasiado en serio su "significado". Estamos siempre conscientes de que nuestros términos son un tanto vagos (puesto que hemos aprendido a darles sólo aplicaciones prácticas), y logramos la precisión, no reduciendo su penumbra de vaguedad, sino conservándolos en esa penumbra, redactando nuestras frases de tal manera que los posibles matices de significado no tengan mucha importancia. Así, evitamos entrar en controversias acerca de las palabras.

El punto de vista de que la precisión de la ciencia y del lenguaje científico depende de la precisión de sus términos es ciertamente muy encomiable, pero en mi opinión no deja de ser un mero prejuicio. La precisión de un lenguaje depende, más bien, de que se tenga el cuidado de no sobrecargarlo con la tarea de ser muy preciso. Términos como "duna" o "viento" son ciertamente muy vagos. (¿Cuántos centímetros o metros debe medir la altura de un montículo de arena para que se le pueda llamar "duna"? ¿Con qué velocidad debe desplazarse una masa de aire para que se le pueda llamar "viento"?). Sin embargo, para muchos usos de los geólogos, estos términos son suficientemente precisos; y para otros usos, cuando se necesita un grado más alto de diferenciación, podemos decir: "dunas de entre 1.20 y 9 metros de altura", o "vientos con velocidad de entre 30 y 60 kilómetros por hora". Y la posición en las ciencias más exactas es análoga. En las mediciones que hacemos en física, por ejemplo, siempre tenemos el cuidado de tomar en cuenta el rango en que puede haber error, o el margen de error, y la precisión no consiste en tratar de reducir este margen a cero, ni en pretender que no hay tal margen de error, sino más bien en reconocerlo explícitamente.

Hasta cuando un término ha suscitado algún problema (por ejemplo, el término "simultaneidad" en física), no ha sido porque su significado resulte ambiguo o impreciso, sino más bien porque alguna teoría intuitiva nos indujo a sobrecargar el término con demasiado significado, o con demasiada "precisión", para no atribuirle muy poca precisión. Lo que Einstein descubrió al analizar la simultaneidad es que, al hablar de acontecimientos simultáneos, los físicos incurrieron en una falsa suposi-

ción, que habría sido inobjetable de haber existido indicios de velocidad infinita. El error no consistía en que no quisieran decir nada con esta expresión, o en que le dieran un significado ambiguo, o en que el término careciera de suficiente precisión; lo que Einstein descubrió fue que la eliminación de una suposición teórica, no advertida hasta entonces por considerarse intuitivamente obvia, podía suprimir una dificultad que había surgido en la ciencia. En concordancia con esto, a él no le preocupaba el significado de un término sino la veracidad de una teoría. Seguramente no se habría llegado a ninguna mejora en este sentido si alguien hubiese empezado, aparte de analizar un problema definido de física, por mejorar el concepto de simultaneidad analizando su "significado esencial", o incluso analizando bien lo que los físicos "realmente quieren decir" cuando hablan de simultaneidad.

Pienso que podemos aprender una lección de este ejemplo: no deberíamos intentar cruzar los puentes antes de llegar a ellos. Y pienso también que la preocupación por cuestiones acerca del significado de los términos, tales como su vaguedad o su ambigüedad, no puede justificarse de ningún modo esgrimiendo el ejemplo de Einstein. Tal preocupación se basa, más bien, en suponer que mucho depende del significado de nuestros términos, y de que operamos según este significado; y, por tanto, esta actitud nos hará caer en el verbalismo y en el escolasticismo. Desde este punto de vista podemos criticar una teoría como la de Wittgenstein,¹⁷ que sostiene que la ciencia investiga los hechos, en tanto

¹⁷ Wittgenstein enseña en el *Tractatus* (véase también la nota 13) que la filosofía no puede hacer proposiciones, y que todas las proposiciones filosóficas son de hecho pseudo-proposiciones sin ningún significado. Estrechamente relacionada con esto está su doctrina de que la verdadera tarea de la filosofía no es proponer enunciados, sino aclararlos: "El objeto de la filosofía es el esclarecimiento lógico de los pensamientos. La filosofía no es una teoría, sino una actividad. El trabajo filosófico consiste esencialmente en hacer elucidaciones" (4.112).

Surge así la pregunta de si este punto de vista está en concordancia con el objetivo fundamental de Wittgenstein, la destrucción de la metafísica quitándole todo valor para revelarla como absurda sin ningún significado. En mi obra *The Logic of Scientific Discovery* [*La lógica de la investigación científica*] (véanse especialmente las pp. 311-314) intenté demostrar que el método de Wittgenstein nos lleva a una solución meramente verbal y que debe de suscitar, pese a su aparente radicalismo, no a la destrucción, ni a la exclusión, ni siquiera a la clara demarcación de la metafísica, sino a su intrusión en el campo de la ciencia, y a su confusión con la ciencia. Las razones para afirmar esto son muy sencillas.

Consideremos uno de los enunciados de Wittgenstein; por ejemplo, "La filosofía no es una teoría, sino una actividad". Por supuesto, no es éste un enunciado que se refiera a "la ciencia natural total" (o a la totalidad de las ciencias naturales). Por tanto, según Wittgenstein, (véase más arriba, la nota 13), no puede pertenecer "a la totalidad de las proposiciones verdaderas". Por otra parte, tampoco es una proposición falsa (ya que, si lo fuera, su negación tendría que ser verdadera, y tendría que pertenecer a las ciencias naturales). Así, llegamos al resultado de que debe ser "sin significado", o "sin ningún sentido", o "absurda"; y lo mismo se aplica a la mayoría de las proposiciones de Wittgenstein. Wittgenstein mismo reconoce esta consecuencia de su doctrina, pues escribe (6.54): "Mis proposiciones son

que la filosofía debe esclarecer el significado de los términos, con lo cual purga nuestro lenguaje y elimina los rompecabezas lingüísticos. Es característica de los puntos de vista de esta escuela no conducir a ningún concatenamiento de argumentos que pudiera criticarse racionalmente; por tanto, la escuela de Wittgenstein dirige sus análisis¹⁸ sutiles exclusivamente al pequeño y esotérico círculo de los iniciados en ella. Esto parece sugerir que cualquier preocupación por el significado tiende a redundar en ese resultado que es tan típico del aristotelismo: el escolasticismo y el misticismo.

elucidatorias de esta manera: quien me entiende finalmente reconoce que carecen de significado..." El resultado es importante. La filosofía de Wittgenstein carece de significado, y él mismo reconoce que así es. Y, como Wittgenstein dice en su Prefacio, "Por otra parte, la veracidad de los pensamientos que aquí se comunican me parece incontrovertible y definitiva. Por tanto, opino que los problemas, en lo esencial, han quedado resueltos". Esto demuestra que podemos comunicar *pensamientos incontrovertibles y definitivamente verdaderos* por medio de proposiciones que se reconocen como sin significado, y que podemos "finalmente" resolver problemas haciendo proposiciones sin sentido alguno. [Véase también la nota 4-2)-b) al escrito 2.]

Reflexionemos en lo que esto significa. Significa, nada menos, que todas las absurdidades metafísicas contra las que han luchado Bacon, Hume, Kant y Russell durante siglos ahora se pueden dirimir cómodamente, y no obstante, reconocerse que todo esto es absurdo. (Heidegger lo hace así.) Porque ahora tenemos a nuestra disposición una nueva clase de absurdidad, la cual comunica pensamientos cuya veracidad es controvertible y definitiva; en otras palabras, se trata de una *absurdidad profundamente significativa*.

No niego que los pensamientos de Wittgenstein sean incontrovertibles y definitivos. Porque, ¿cómo podría uno atacarlos? Obviamente, cualquier cosa que uno diga en contra de ellos debe ser algo filosófico y, por tanto, absurdo. Y es posible desechar estos ataques como absurdos. Así nos enfrentamos a una posición que en otro lugar he descrito, en relación con Hegel (véase *Conjectures and Refutations* [*Conjeturas y refutaciones*], p. 327 [y también, más adelante, la p. 396 y s], como *dogmatismo reforzado*. [Véase también, más adelante, la pp. 150 y s.]

Resumamos: la teoría antimetafísica del significado, en el *Tractatus* de Wittgenstein, lejos de contribuir a combatir el dogmatismo metafísico y la filosofía oracular, representa un dogmatismo reforzado que abre la puerta de par en par al enemigo, la absurdidad metafísica de significado profundo, y que expulsa, por la misma puerta, a nuestra mejor amiga, es decir, a la hipótesis científica. [Véase también el escrito 8, acerca del *problema de la demarcación*.]

¹⁸ Parece así que el irracionalismo, en el sentido de ser una doctrina o un credo que no propone argumentos articulados y debatibles, sino aforismos y afirmaciones dogmáticas que deben "entenderse" o dejarse a un lado, tenderá generalmente a convertirse en la propiedad de un círculo esotérico de iniciados. Y, ciertamente, algunas publicaciones que se basan en la escuela de Wittgenstein parecen corroborar este pronóstico. (No deseo generalizar; por ejemplo, todo cuanto he leído de los escritos de F. Waismann está presentado como un encadenamiento de argumentos racionales y clarísimos, y del todo libre de esa actitud de "tómalo o déjalo".)

Algunas de esas publicaciones esotéricas parecen no plantear ningún problema serio; para mí, parecen sutiles en obsequio de la sutileza misma. Es significativo que procedan de una escuela que empezó por denunciar a la filosofía por la vaga sutileza de sus conceptos para ocuparse de los seudoproblemas.

Puedo terminar esta crítica al declarar brevemente que no creo que haya mucha justificación para luchar en contra de la metafísica en general, o que pueda resultar algo valioso de esa lucha. Es necesario resolver el problema de la demarcación entre la ciencia

Consideremos brevemente cómo han surgido estos dos resultados típicos del aristotelismo. Aristóteles insistía en que la demostración o comprobación y la definición son los dos métodos fundamentales para obtener el conocimiento. Si analizamos primero la doctrina de la comprobación, no podemos negar que ha conducido a incontables intentos de probar más de lo que puede probarse; la filosofía medieval está llena de este escolasticismo, y la misma tendencia puede observarse, en Europa, hasta el advenimiento de Kant. Fue la crítica de Kant, de todos los intentos para probar la existencia de Dios, la que desembocó en la reacción romántica de Fichte, Schelling y Hegel. La nueva tendencia consiste en descartar las pruebas y, con ellas, cualquier clase de argumento racional. Con los románticos, se pone de moda una nueva clase de dogmatismo, tanto en filosofía como en las ciencias sociales. Nos enfrenta a su *dictum*, y podemos *tomarlo o dejarlo*. Este periodo romántico de una filosofía oracular, al que Schopenhauer llamó la "era de la falta de honradez", lo describe este autor así:¹⁹

El carácter de la honradez, de la probidad intelectual, ese espíritu de emprender una búsqueda en compañía del lector, que impregna las obras de todos los filósofos anteriores, desaparece completamente aquí. Cada página da testimonio de que estos llamados "filósofos" no se proponen enseñar, sino embujar al lector.

Un resultado similar a éste tuvo la doctrina aristotélica de la definición. Primero dio lugar a muchísimas sutilezas. Pero después los filósofos empezaron a sentir que no es posible argumentar acerca de las definiciones. De esta manera, el esencialismo no sólo alentó el verbalismo, sino que originó una desilusión respecto a los argumentos; es decir, una desilusión

y la metafísica. Pero deberíamos reconocer que muchos sistemas metafísicos han redundado en importantes resultados científicos. Me referiré sólo, en este sentido, al sistema de Demócrito; y al de Schopenhauer, que es muy similar al de Freud. Y algunos sistemas metafísicos, por ejemplo, los de Platón, o Malebranche o Schopenhauer, son bellas estructuras de pensamiento. Pero, al mismo tiempo, creo que deberíamos luchar contra aquellos sistemas metafísicos que tienden a hechizarnos y a confundirnos. Pero debe quedar claro que deberíamos hacer lo mismo incluso con los sistemas no metafísicos y antimetafísicos, si manifiestan esta peligrosa tendencia. Y creo que no podemos hacerlo de un solo impulso. Más bien debemos tomarnos el trabajo de analizar los sistemas con cierto detalle; debemos demostrar que entendemos lo que el autor quiere decir, pero que lo que quiso decir no merece la pena de entenderlo. (Es característico de todos estos sistemas dogmáticos, y especialmente de los sistemas esotéricos, que sus admiradores aseveren de todos sus críticos que "no entienden"; pero los tales admiradores olvidan que la comprensión, el entendimiento cabal de algo, debe conducirnos al consenso únicamente en el caso de enunciados con un contenido trivial. En todos los demás casos, podemos entender y disentir.)

¹⁹ Véase A. Schopenhauer, *Die beiden Grundprobleme der Ethik* [Fundamentos de la moral], 4ª ed., 1890, p. 147.

de la razón. El escolasticismo, el misticismo y el desesperar de la razón: he ahí los inevitables resultados del esencialismo de Platón y de Aristóteles. Y la abierta revuelta de Platón en contra de la libertad se convierte, en Aristóteles, en una secreta revuelta contra la razón.

Como sabemos por palabras de Aristóteles mismo, el esencialismo y la teoría de la definición chocaron con fuerte oposición cuando fueron propuestos, especialmente por parte de Antístenes, viejo compañero de Sócrates cuya crítica parece haber sido muy sensata.²⁰ Pero, desafortunadamente, esta oposición fue derrotada. Las consecuencias de esta derrota para el desarrollo intelectual de la humanidad fueron de un alcance incalculable.

²⁰ Simplicio, una de nuestras mejores fuentes sobre estos dudosos asuntos, presenta a Antístenes (*ad Arist. Categ.* 66b, 67b) como un oponente a la teoría de Platón de las Formas o Ideas, y, de hecho, a la doctrina del esencialismo y de la intuición intelectual. Se dice que Antístenes dijo: "Puedo ver un caballo, Platón, pero no puedo ver su caballidad". (Un argumento muy parecido a éste se atribuye a una fuente menor, Diógenes Laercio, vi, 53, dirigido a Diógenes el Cínico, y no hay razón para que este último no lo haya utilizado también.) Pienso que podemos confiar en Simplicio (que al parecer tuvo acceso a Teofrasto), considerando que el propio testimonio de Aristóteles, en la *Metafísica* (especialmente en 1043b24) se compagina bien con este antiesencialismo de Antístenes.

Deseo agregar aquí que, pese a todas mis críticas, estoy más que dispuesto a aceptar los méritos de Aristóteles. Es el fundador de la lógica, y hasta llegar a los *Principia Mathematica*, podemos afirmar que toda la lógica es una elaboración y generalización de los comienzos aristotélicos. (No cabe duda de que ha comenzado una nueva época en lógica, en mi opinión, aunque no con la llamada "lógica no aristotélica" o con los sistemas "multivalentes", sino más bien con la clara distinción entre "lenguaje objeto" y "metalinguaje".) Además, Aristóteles tiene el gran mérito de haber intentado la doma del idealismo con su enfoque de sentido común, que insiste en que sólo las cosas individuales son "reales" (y en que sus "formas" y "materia" son aspectos o abstracciones). Sin embargo, este mismo enfoque es responsable de que Aristóteles mismo no intente siquiera resolver el problema platónico de los universales [véanse las pp. 182 y 183 más arriba], es decir, el problema de explicar por qué ciertas cosas se parecen unas a otras y otras no. Pues, ¿por qué no habría tantas diferentes esencias aristotélicas en las cosas, tantas como cosas existentes?

7. EL PROBLEMA DE LA INDUCCIÓN (1953, 1974)

PARA formular brevemente el problema de la inducción, podemos referirnos a Born, quien escribe: "...ninguna observación, ningún experimento, por extensos que sean, pueden dar más que un número finito de repeticiones"; por tanto, "el enunciado de una ley —B depende de A— trasciende siempre al experimento. No obstante, esta clase de enunciado se hace por todas partes y todo el tiempo, y a veces se basa en muy escaso material".¹

En otras palabras, el problema lógico de la inducción surge de: 1) el descubrimiento de Hume (que tan bien expresó Born) de que es imposible justificar una ley mediante la observación o la experimentación, ya que "trasciende la experimentación"; 2) el hecho de que la ciencia propone y utiliza leyes "por todas partes y todo el tiempo". (Como a Hume, a Born le impresiona "el escaso material", es decir, los pocos casos observados en que la ley puede fundamentarse.) A esto podemos agregar: 3) *el principio del empirismo*, que asevera que, en la ciencia, sólo la observación y la experimentación pueden decidir sobre la *aceptación* o el *rechazo* de los enunciados científicos, incluyendo las leyes y las teorías.

Estos tres principios, 1), 2) y 3), parecen a primera vista chocar unos con otros; y este choque aparente constituye el *problema lógico de la inducción*.

Enfrentado a este choque, Born desecha el principio 3), el del empirismo (como Kant y otros muchos, incluso Bertrand Russell, ya lo habían hecho antes que él), en favor de lo que él llama el "principio metafísico"; un principio metafísico que ni siquiera intenta formular, que describe vagamente como un "código o regla de oficio", y del que no he visto nunca ninguna formulación que al menos pareciera prometedora y no fuera claramente insostenible.

Pero, en realidad, los principios 1) a 3) no se oponen entre sí. Podemos ver esto en el momento en que advertimos que la aceptación por parte de la ciencia de una ley o de una teoría es *sólo tentativa*; lo cual equivale a decir que todas las leyes y teorías son conjeturas, o *hipótesis*

¹ Véase M. Born, *Natural Philosophy of Cause and Chance* [La filosofía natural de la causa y del azar], 1949, p. 6.

tentativas (posición que a veces he llamado "*hipoteticismo*"); y que podemos rechazar una ley o una teoría a la luz de nuevas evidencias, sin que por ello tengamos que desechar necesariamente la vieja evidencia que nos hizo aceptarla originalmente. (No dudo que Born y otros muchos estarían de acuerdo en que las teorías se aceptan sólo tentativamente; pero la muy difundida creencia en la inducción demuestra que rara vez se advierten las implicaciones de largo alcance de este punto de vista.)

El principio 3) del empirismo puede mantenerse plenamente, puesto que el destino de una teoría, su aceptación o su rechazo lo deciden la observación y la experimentación... mediante el resultado de las pruebas. En la medida en que una teoría soporta las más rigurosas pruebas que podamos diseñar, es aceptada; en el caso contrario, es rechazada. Pero nunca se infiere, en ningún sentido, de la evidencia empírica. No existe la inducción, ni psicológica ni lógica. *Sólo la falsedad de una teoría puede inferirse de la evidencia empírica, y esta inferencia es puramente deductiva.*

Hume demostró que no es posible inferir una teoría de afirmaciones de observación; pero esto deja en pie la posibilidad de refutar una teoría por afirmaciones de observación; la apreciación irrestricta de esta posibilidad esclarece perfectamente la relación entre teorías y observaciones.

Así se resuelve el problema del pretendido choque entre los principios 1), 2) y 3), y también el problema de la inducción que planteó Hume.

El problema de la inducción de Hume casi siempre ha sido mal formulado por lo que podemos llamar la tradición filosófica. Daré primero unos cuantos ejemplos de estas malas formulaciones, que llamaré las *formulaciones tradicionales del problema de la inducción*. Pero las sustituiré con lo que yo considero mejores formulaciones.

Típicos ejemplos de formulaciones del problema de la inducción que son a la vez malas y tradicionales son los siguientes.

¿Cuál es la justificación de la creencia de que el futuro se asemejará al pasado? ¿Cuál es la justificación de las llamadas *inferencias inductivas*?

Se entiende aquí por inferencia inductiva una inferencia de *casos observados* repetidamente para concluir de ellos *casos aún no observados*. Tiene una importancia comparativamente menor si tal inferencia de lo observado hacia lo no observado es, desde el punto de vista del tiempo, predictiva o retrodictiva; si inferimos que el sol saldrá mañana

o que salió hace 100000 años. Por supuesto, desde un enfoque pragmático, podríamos decir que el tipo predictivo de inferencia es el más importante. Y en efecto, así es, sin duda.

Otros varios filósofos también consideran mal concebido este tradicional problema de la inducción. Algunos afirman que está mal concebido porque no se necesita dar ninguna justificación a la inferencia inductiva; y de hecho, tampoco es necesaria para la inferencia deductiva. La inferencia inductiva es inductivamente válida, así como la inferencia deductiva es deductivamente válida. Creo que fue el profesor Strawson el primero que dijo esto.

Soy de distinta opinión. Sostengo, siguiendo a Hume, que simplemente no existe una entidad lógica considerada inferencia inductiva; o sea, que todas las llamadas inferencias inductivas son lógicamente inválidas... incluso *inductivamente* inválidas, para decirlo más rotundamente [véase el final de este escrito]. Tenemos muchos ejemplos de inferencias deductivamente válidas, e incluso algunos criterios parciales de validez deductiva; pero no existe ningún ejemplo de una inferencia inductivamente válida.² Y sostengo, dicho sea de paso, que este resultado puede encontrarse en la obra de Hume, aunque Hume, al mismo tiempo, y en radical contraste conmigo, *creía en el poder psicológico de la inducción*; no como un procedimiento válido, sino como un procedimiento que los animales y los hombres utilizan con éxito, como algo real y de necesidad biológica.

Considero tarea importante aclarar, aun a costa de ser algo repetitivo, en qué estoy de acuerdo con Hume y en qué discrepo de él.

Estoy de acuerdo con la opinión de Hume de que la inducción es inválida y que en ningún sentido se justifica. Por consiguiente, ni Hume ni yo podemos aceptar las formulaciones tradicionales que, sin estar de por medio la crítica, piden la justificación de una inducción; tal petición es acrítica, porque no ve la posibilidad de que la inducción sea inválida *en todos los sentidos*, y por tanto *injustificable*.

Discrepo de la opinión de Hume (de paso, la opinión de casi todos los filósofos) de que la inducción es un hecho y, en todo caso, una necesidad. Sostengo que ni los animales ni los hombres utilizan ningún procedimiento como la inducción, ni ningún argumento basado en la repetición de ejemplos. La creencia de que utilizamos la inducción es simplemente una equivocación. Es una especie de ilusión óptica.

Lo que sí utilizamos es un método de prueba y eliminación de error; por más que engañosamente se confunda este método con la inducción, si

² Véanse, por ejemplo, las secciones 10 y 11 de mi comentario sobre la obra de Carnap, en las pp. 285-303 de I. Lakatos (comp.), *The Problem of Inductive Logic* [El problema de la lógica inductiva], 1968, y la sección 32 de *Búsqueda sin término*.

lo examinamos bien, su estructura lógica difiere totalmente de la estructura lógica de la inducción. Además, es un método que no hace surgir ninguna de las dificultades relacionadas con el problema de la inducción.

Por tanto, no es porque la inducción pueda manejarse sin justificación por lo que me opongo al tradicional problema; al contrario, yo necesitaría urgentemente que hubiera justificación; pero esta necesidad no puede satisfacerse. Porque, sencillamente, la inducción no existe, y el punto de vista contrario es una garrafal equivocación.

ni

Hay muchas maneras de presentar mi propia opinión, no inductivista. Quizá la más sencilla sea ésta: trataré de demostrar que todo el aparato de la inducción se vuelve innecesario en cuanto aceptamos la falibilidad general del conocimiento humano o, como me gusta llamarla, el *carácter conjetural del conocimiento humano*.

Permítaseme señalar primero lo que pienso de la mejor clase de conocimiento que tenemos: esto es, del conocimiento científico. Afirmando que el conocimiento científico es, esencialmente, conjetural o hipotético.

Tomemos como ejemplo la mecánica clásica newtoniana. Jamás hubo una teoría de tanto éxito como ésta. Si el éxito repetido de las observaciones podía establecer la validez de una teoría, tal habría sido la de Newton. Sin embargo, la teoría de Newton fue superada en el campo de la astronomía por la teoría de Einstein, y en el campo atómico por la teoría cuántica. Y casi todos los físicos, hoy, piensan que la teoría clásica de la mecánica newtoniana no es sino una maravillosa conjetura, una hipótesis que tuvo un extraño éxito y una asombrosamente buena aproximación a la verdad.

Ahora puedo formular mi tesis central, a saber: en cuanto nos damos cuenta cabal de las implicaciones de la índole conjetural del conocimiento humano, el problema de la inducción cambia radicalmente de aspecto: no hay ya necesidad de que nos inquieten los resultados negativos de Hume, puesto que no existe ya la necesidad de atribuir al conocimiento humano una *validez* derivada de repetidas observaciones. El conocimiento humano carece de tal validez. Por otra parte, podemos explicar todos nuestros logros en términos del método de prueba y eliminación de error. Para decirlo en pocas palabras, nuestras conjeturas son nuestros globos-sonda, y las probamos criticándolas y tratando de sustituirlas, tratando de demostrar que puede haber mejores o peores conjeturas, y que pueden mejorarse. El lugar del problema de la

inducción lo ocupa el problema de la comparativa bondad o imperfección de las conjeturas o teorías rivales que se hayan propuesto.

La barrera más importante para que se acepte la índole conjetural del conocimiento humano, y para aceptar que contiene en sí misma la solución del problema de la inducción, es una doctrina a la que podemos llamar la teoría de sentido común sobre el conocimiento humano, o *teoría de la cubeta de la mente humana*.³

IV

Tengo en gran estimación el sentido común. En realidad, pienso que toda la filosofía debe comenzar desde puntos de vista del sentido común y desde su examen crítico.

Para nuestros propósitos, deseo en este punto distinguir dos partes del enfoque de sentido común acerca del mundo, y llamar la atención sobre el hecho de que chocan entre sí.

La primera parte se refiere al realismo de sentido común; consiste en que hay un mundo real, con gente real, animales y plantas reales, autos y estrellas reales en él. Creo que este enfoque es verdadero e importantísimo, y que jamás se ha propuesto acerca de él ninguna crítica válida. [Véase, más adelante, el escrito 17.]

Otra parte muy diferente del enfoque de sentido común acerca del mundo es la *teoría del conocimiento* de sentido común. El problema consiste en cómo obtenemos nuestro conocimiento acerca del mundo. La "solución" de sentido común es ésta: abrir bien los ojos y los oídos. *Nuestros sentidos son las principales, si no las únicas, fuentes de nuestro conocimiento del mundo.*

Considero este segundo enfoque tan del todo erróneo como insuficientemente criticado (pese a Leibniz y pese a Kant). Llamo a este enfoque la "teoría de la cubeta de la mente", porque puede resumirse con el diagrama siguiente:



³ Véanse también el apéndice 1 y el capítulo 2 de *Conocimiento objetivo*.

Lo que se supone entra en la cubeta a través de nuestros sentidos son los elementos, los átomos o moléculas, del conocimiento. Por consiguiente, nuestro conocimiento consiste en una acumulación, un compendio, o quizás una síntesis de los elementos que nos ofrecen nuestros sentidos.

Ambas mitades de la filosofía del sentido común, el realismo de sentido común y la teoría de sentido común del conocimiento, fueron sostenidas por Hume; él descubrió, como lo había hecho Berkeley antes, que hay un choque entre estos dos enfoques. Porque la teoría de sentido común del conocimiento puede conducir hacia una especie de antirrealismo. Si el conocimiento resulta de sensaciones, entonces las sensaciones son los únicos elementos *ciertos* (veraces) de conocimiento, y no nos asiste ninguna buena razón para creer que exista nada fuera de nuestras sensaciones.

Hume, Berkeley y Leibniz creían en el principio de la razón suficiente. Para Berkeley y para Hume, tomaba esta forma: si no tienes suficientes razones para creer algo, entonces este hecho es razón suficiente para abandonar tal creencia. Tanto para Berkeley como para Hume, el conocimiento auténtico consistía esencialmente en creencias, apoyadas en razones suficientes: pero esto les hizo tomar la posición de que el conocimiento consiste, más o menos, en ciertas sensaciones subjetivas.

Así pues, para estos filósofos, el mundo real de sentido común no existe en realidad; según Hume, ni siquiera nosotros mismos existimos plenamente. Lo único que existe son sensaciones, impresiones, imágenes de la memoria. [Véase también el escrito XXII, sección i, más adelante.]

Esta visión antirrealística puede caracterizarse con varios nombres, pero el más usual parece ser "idealismo". El idealismo de Hume fue en su concepto una refutación estricta del realismo de sentido común. Pero, aunque se sentía *racionalmente* obligado a mirar como un error el realismo de sentido común, él mismo reconocía que en la práctica le era imposible dejar de creer en el realismo de sentido común durante más de una hora.

Así, Hume tuvo la fuerte experiencia de observar el choque entre las dos partes de la filosofía de sentido común: el realismo y la teoría del conocimiento basada en el sentido común. Y aunque se daba cuenta de que emocionalmente no podía prescindir del realismo consideró que este hecho era una mera consecuencia de la costumbre irracional o del hábito; estaba convencido de que la constante adherencia a los resultados más críticos de la teoría del conocimiento nos haría abandonar el realismo.⁴ Fundamentalmente, el idealismo de Hume sigue siendo la corriente principal del empirismo inglés.

⁴ D. Hume, *A Treatise of Human Nature [Tratado sobre la naturaleza humana]*, Libro i, Cuarta Parte, el último párrafo de la sección II; edición de L. Selby-Bigge, p. 218. David

Los dos problemas de Hume relacionados con la inducción —el problema lógico y el problema psicológico— pueden exponerse mejor, en mi opinión, con el trasfondo de la teoría de la inducción basada en el sentido común. Esta teoría es muy sencilla: como se supone que todo el conocimiento es resultado de pasadas observaciones, así se espera especialmente que sea todo conocimiento que se da por sentado, como que el sol saldrá mañana, que todos los hombres morirán algún día o que el pan alimenta. Todo esto tiene que ser resultado de observaciones hechas en el pasado.

El mérito inmortal de Hume es que se atrevió a poner en tela de juicio el enfoque de la inducción basada en el sentido común, aunque nunca dudó de que debía de ser verdadero en gran medida. Creía que la inducción por repetición era lógicamente insostenible; que racional o lógicamente *ninguna cantidad* de instancias observadas puede tener el menor peso en las instancias no observadas. Esto representa la solución negativa de Hume al problema de la inducción, solución a la que me adhiero sin reservas.

Pero, al mismo tiempo, Hume sostenía que, si bien la inducción era racionalmente inválida, era un hecho psicológico, y todos deberíamos confiar en ella.

Por tanto, los dos problemas de Hume relacionados con la inducción eran los siguientes:

1) El problema lógico: *¿Estamos justificados racionalmente al razonar a partir de repetidas instancias de las que tuvimos experiencia, para llegar a instancias de las que no hemos tenido ninguna experiencia?*

La inquebrantable respuesta de Hume fue: No; no estamos justificados para ello, por más grande que sea el número de instancias repetidas. Y añadía que para él no había la menor diferencia si, al tratar este problema, pedimos la justificación, no de *cierta* creencia, sino de una creencia *probable*. Las instancias de las que hemos tenido experiencia no nos permiten razonar ni argumentar acerca de la *probabilidad* de instancias de las que no hayamos tenido experiencia, ni tampoco acerca de la *certeza* o certidumbre de tales instancias.

Miller me ha señalado que Hume, al establecer y experimentar el contraste entre lo que creía (el realismo) y lo que pensaba que era verdadero (idealismo), dio con esto —sin duda involuntariamente— un primer paso para alejarse de su propia caracterización (de sentido común) del conocimiento como una forma de la creencia; es decir, un paso hacia el reconocimiento del profundo abismo que hay entre el mundo número 2 y el mundo número 3. Por desgracia, este descubrimiento de Hume se quedó en el nivel del mundo número 2, y no logró convertirse en un problema objetivo en el nivel del mundo número 3.

2) La pregunta psicológica siguiente: *¿Cómo es que, aunque todas las personas razonables esperan y creen que las instancias de las que no han tenido experiencia, se conformarán a aquellas de las que sí han tenido experiencia?* O, en otras palabras, ¿por qué todos tenemos *expectativas*, y por qué nos aferramos a ellas con tan gran *confianza* o con tan firme creencia?

La respuesta de Hume a este problema psicológico de la inducción fue: *"Por costumbre o hábito"*; o, en otras palabras, *por el irracional pero irresistible poder de la ley de asociación*. Estamos condicionados por la repetición; es un mecanismo de acondicionamiento sin el cual, dice Hume, apenas podríamos sobrevivir.

Mi opinión es que la respuesta de Hume al problema lógico es la correcta, y que su respuesta al psicológico es un error, pese a su gran poder de persuasión.

vi

Las respuestas de Hume a los problemas lógico y psicológico de la inducción nos llevan de inmediato a una conclusión irracionalista. Según Hume, todo nuestro conocimiento, especialmente nuestro conocimiento científico, es sólo hábito o costumbre irracional, y racionalmente es totalmente indefendible.

Hume mismo consideraba esto una especie de escepticismo; pero era más bien, como lo señaló Bertrand Russell, un sometimiento inintencional al irracionalismo. Es asombroso que un genio sin igual de la crítica, una de las mentes más racionalistas de todos los tiempos, no sólo llegara a no creer en la razón, sino que se volviera un campeón de la sinrazón, del irracionalismo.

Nadie ha advertido tan bien esta paradoja como Bertrand Russell, admirador y, en muchos aspectos, tardío discípulo de Hume. Así, en el capítulo dedicado a Hume de su *Historia de la filosofía occidental*, publicada en 1946, Russell dice acerca del tratamiento que da Hume a la inducción: "La filosofía de Hume [...] representa la bancarrota de la razonabilidad del siglo xvm", y agrega: "Por ello, es importante descubrir si hay alguna respuesta a Hume dentro de la filosofía que es total o principalmente *empírica*. De no ser así, *no existe ninguna diferencia intelectual entre la cordura y la locura*. Al lunático que cree que él es un huevo escalfado, sólo se le condenará con fundamento en que pertenece a una minoría..."

Russell prosigue aseverando que, si la inducción (o el principio de inducción) es rechazada, "cada intento de llegar a formular leyes científicas generales a partir de observaciones particulares resulta falaz, y el empirista no podrá escapar del escepticismo de Hume".

Y Russell resume su opinión acerca de la situación que suscitó el choque entre las dos respuestas de Hume con esta dramática observación:

*"El desarrollo de la sinrazón durante todo el siglo xix en lo que va del xx, es la secuela natural de la destrucción del empirismo, que llevó a cabo Hume."*⁵

Esta última cita de Russell *acaso* exagere un poco. No deseo dramatizar demasiado la situación; y aunque a veces siento **que Russell está en lo cierto** al hacer hincapié en esto, en otros momentos **lo dudo**.

Sin embargo, la siguiente cita del profesor Strawson **me parece que** apoya la grave opinión de Russell: "[Si] ...existe el problema **de la inducción** y [...] Hume lo planteó, debe añadirse que él lo resolvió... [:] nuestra aceptación de los 'cánones básicos' [de la inducción] [...] **nos la impone la naturaleza** [...] La razón es, y debe ser, **la esclava de las pasiones**".⁶

Sea de ello lo que fuere, afirmo que tengo una respuesta **para el** problema psicológico de Hume que anula por completo el **choque entre la** lógica y la psicología del conocimiento; y que, además, **nulifica todos** los razonamientos de Hume y de Strawson en contra de la **razón**.

vn

Mi propia manera de soslayar las consecuencias **irracionalistas de** Hume es muy sencilla. Resuelvo el problema psicológico **de la inducción** (y también las formulaciones tales como el problema **pragmático**) de una manera que satisface el siguiente *"principio de la primacía de la solución lógica"* o, más brevemente, el *"principio de transferencia"*. Este principio puede expresarse así: la solución del **problema lógico de la** inducción, lejos de chocar con las de los problemas **psicológico o pragmático**, puede, procediendo con cierto cuidado, **transferirse** directamente a esas soluciones. Así, no habrá choque, y **no hay consecuencias** irracionalistas.

Para empezar, el problema lógico de la inducción **necesita una** reformulación.

⁵ Las tres citas son de Bertrand Russell, *A History of Western Philosophy [Historia de la filosofía occidental]*, 1946, pp. 698 y ss; nueva edición, 1961, pp. 645-647. (Las *italicas* son mías.)

⁶ Véase la p. 21 de P. F. Strawson, "On Justifying Induction" ["**De la justificación de la inducción**"], en *Philosophical Studies* 9, 1958, pp. 20 y ss. Véase también Hume, *op. cit.*, Libro II, Tercera Parte, sección m; edición de L. Selby-Bigge, p. 415: "**La razón es, y sólo debe ser, la esclava de las pasiones...**"

Primero, debe formularse en términos, no sólo de "instancias" (como lo formuló Hume), sino en términos de regularidades universales o leyes. Las regularidades o leyes están presupuestas en el término de Hume "instancia"; porque una instancia es una instancia *de* algo..., de una regularidad o de una ley. (O, mejor dicho, es una instancia de muchas regularidades o de muchas leyes.)

En segundo lugar, debemos ampliar el alcance de los razonamientos de instancias a leyes para que también podamos tomar en cuenta las **contrainstancias**.

De esta manera, llegamos a la reformulación del *problema lógico de la inducción* de Hume en los términos siguientes:

¿Estamos racionalmente justificados al razonar a partir de instancias o de contrainstancias de las que hemos tenido alguna experiencia, para descubrir la veracidad o falsedad de las correspondientes leyes, o llegar a instancias de las que antes no hemos tenido ninguna experiencia?

Esto constituye un problema puramente lógico. Es esencialmente sólo una pequeña extensión del problema lógico de la inducción de Hume formulado aquí antes, en la sección v.

La respuesta a este problema es como sigue: como está implícito en el texto de Hume, en verdad no estamos justificados al razonar a partir de una instancia hacia la búsqueda de la veracidad de la ley correspondiente. Pero de esta negativa se colige un segundo resultado, también negativo, por lo cual podemos agregar: *sí* estamos justificados al razonar a partir de una contrainstancia en cuanto a la *falsedad* de la correspondiente ley universal (es decir, de cualquier ley de la que sea una contrainstancia). O, en otras palabras, a partir de un punto de vista puramente lógico, aceptar *una* contrainstancia a la aseveración "Todos los cisnes son blancos" implica la falsedad de la ley "Todos los cisnes son blancos", es decir, de esa ley cuya contrainstancia hemos aceptado. La inducción es lógicamente inválida; pero la refutación o falsificación es un modo lógicamente válido de argumentar a partir de una contrainstancia en cuanto a (o, más bien, contra) la ley correspondiente.

Esto demuestra que sigo estando de acuerdo con Hume en su punto de vista negativo en cuanto al resultado lógico; pero de esta manera lo he ampliado.

Esta situación lógica es del todo independiente de cualquier pregunta o de si podríamos, en la práctica, aceptar una sola contrainstancia —por ejemplo, un solitario cisne negro— para refutar una ley que haya tenido gran éxito. No sugiero que nos deberíamos dar por satisfechos tan fácilmente; podríamos sospechar que el espécimen negro que está ante nosotros no es un cisne. Y de cualquier modo, en la práctica, nos mostraríamos renuentes a aceptar una contrainstancia aislada. Pero esto es

otra cuestión [véase la sección rv del escrito 10, más adelante]. La lógica nos obliga a rechazar hasta la ley que haya tenido más éxito en el momento en que aceptamos una sola contrainstancia.

Por tanto, podemos decir: Hume estaba en lo correcto en cuanto a su resultado negativo, en el sentido de que no puede haber ningún argumento positivo, lógicamente válido, que nos lleve en la dirección inductiva. Pero existe otro resultado negativo; hay argumentos negativos lógicamente válidos que nos orientan hacia una dirección inductiva: *una contrainstancia puede invalidar o desaprobar una ley*.

VIII

El resultado negativo de Hume establece definitivamente que todas nuestras leyes o teorías universales siguen siendo por siempre tanteos, conjeturas, hipótesis. Pero el segundo resultado negativo respecto de la fuerza de las contrainstancias de ninguna manera descarta la posibilidad de formular una teoría positiva acerca de cómo, sólo con argumentaciones racionales, podemos *preferir* algunas conjeturas que compitan con otras.

De hecho, podemos así construir una muy elaborada *teoría lógica de la preferencia*; preferencia desde el punto de vista de la búsqueda de la verdad.

Para decirlo abreviadamente, la enérgica observación de Russell de que, si rechazamos —siguiendo a Hume— toda inducción positiva, "no hay diferencia intelectual entre la cordura y la locura", es errónea. Porque rechazar la inducción no nos impide preferir, por ejemplo, la teoría de Newton a la de Kepler, o la teoría de Einstein a la de Newton: en nuestra discusión racionalista crítica de estas teorías, *quizás* hayamos aceptado la existencia de contraejemplos respecto de la teoría de Kepler que no refutan la de Newton, y contraejemplos respecto de la teoría de Newton que no refutan la de Einstein. Dada la aceptación de estos contraejemplos, podemos decir que las teorías de Kepler y las de Newton son sin lugar a duda falsas; mientras que la de Einstein puede ser verdadera o falsa: eso no lo sabemos. Así, puede haber preferencias *puramente intelectuales* para una u otra de estas teorías; y estamos muy lejos de poder afirmar, siguiendo a Russell, que desaparece la única diferencia entre la ciencia y la insania. Es cierto que el argumento de Hume queda en pie y que, por tanto, la diferencia entre un científico y un lunático no es que el primero base sus teorías con toda seguridad en la observación, mientras que el segundo no, o algo por el estilo. No obstante, ahora podemos advertir que *puede haber una diferencia* en esto: *acaso*

la teoría del lunático sea fácilmente refutable mediante la observación, mientras que la teoría del científico haya soportado rigurosas pruebas o comprobaciones.

Lo que tienen en común las teorías del científico y las del lunático es que ambas pertenecen al *conocimiento conjetural*. Pero algunas conjeturas son mucho mejores que otras; y ésta es una respuesta suficiente para Russell, y basta para apartarnos del escepticismo radical. Porque, ya que es posible que algunas *conjeturas* sean preferibles a otras, también es posible que nuestro conocimiento *conjetural* mejore, y que *crezca*. (Por supuesto, es posible que una teoría que sea preferida a otra en determinada época pierda nuestra preferencia posteriormente, de manera que la otra sea hoy la preferida. Pero, por otra parte, también es posible que esto no suceda.)

Acaso *prefiramos* algunas teorías en detrimento de otras con fundamento sólo en el racionalismo. Por ello, es importante esclarecer perfectamente cuáles son los principios de preferencia o selección, en este aspecto.

Ante todo, estos principios están regidos por la idea de la verdad. Queremos, hasta donde sea posible, teorías verdaderas, verídicas, y por esta razón intentamos eliminar las falsas.

Pero queremos algo más: queremos descubrir una verdad más nueva y más interesante. Así llegamos a la idea del *crecimiento de los contenidos informativos* y, especialmente, del *contenido de la verdad*, del contenido de veracidad. Esto es, llegamos así al siguiente *principio de preferencia*: en general, una teoría de gran contenido informativo es más interesante, incluso antes de que se haya puesto a prueba, que una teoría con escaso contenido informativo. Ciertamente es que acaso tengamos que desechar la teoría con mayor contenido, o la teoría más osada, como la llamo yo, si no soporta todo el peso de las pruebas. Pero aun en este caso quizá hayamos aprendido más de ella que de una teoría con escaso contenido informativo, pues las pruebas de "falsificación" (para encontrar las falsedades) pueden revelar, a veces, nuevos e inesperados hechos y problemas. [Véase también, más adelante, el escrito 13.]

Así, nuestro análisis lógico nos lleva directamente a una teoría de método, y especialmente a la siguiente regla metodológica: intenta, y apunta hacia, teorías más osadas, más llenas de información, con gran contenido informativo; y luego deja que estas teorías de muchos datos compitan entre sí, discutiéndolas con espíritu crítico y probándolas con mucho rigor.

DC

Mi *solución* al problema lógico de la inducción era que podemos tener ciertas *preferencias* por determinadas conjeturas en competición; esto es, por aquellas que son muy informativas y que hasta ese momento han soportado el rigor de la crítica eliminatoria. Estas conjeturas preferidas son resultado de la selección, de la lucha por sobrevivir de hipótesis sometidas a la presión de *la crítica*, que equivale a *presión de selección artificialmente intensificada*.

Lo mismo se aplica al problema psicológico de la inducción. También aquí nos enfrentamos a hipótesis en competición, que acaso podamos llamar creencias, y algunas de ellas son eliminadas, mientras que otras sobreviven, al menos durante un tiempo. Los animales son con frecuencia eliminados, junto con sus creencias; o bien sobreviven con ellas. Los hombres con frecuencia viven más que sus creencias; pero tanto tiempo como sobrevivan las creencias (a menudo, muy poco tiempo), forman *la base de la acción* (momentánea o durable).

Mi tesis es que este procedimiento darwiniano de la selección de las creencias y acciones no puede describirse, en ningún sentido, como irracional. En ningún modo choca con la solución racionalista del problema lógico de la inducción. Más bien es sólo la transferencia de la solución lógica al campo psicológico. (Lo cual no significa, por supuesto, que jamás sufram de lo que se llama "creencias irracionales".)

Así, con una aplicación del principio de transferencia al problema psicológico de Hume, desaparecen las conclusiones irracionalistas de Hume.

Al hablar de preferencia, hasta aquí sólo he analizado las preferencias del teórico... si acaso tiene alguna; y por qué se inclinará a preferir "la mejor", es decir, la teoría más comprobable, y a preferir la probada de mejor forma. Claro está que acaso el teórico no tenga *ninguna* preferencia; quizá lo desanime la solución "escéptica" de Hume (y la mía) al problema lógico de Hume; acaso se diga que, si no puede *estar seguro* de encontrar la teoría verídica entre las teorías en competición, no le interesa aplicar cualquier método como el descrito, ni siquiera si tal método le da una razonable seguridad de que, *si* una teoría verídica se encontrara entre las teorías propuestas, estará entre las sobrevivientes, las preferidas, las corroboradas. Sin embargo, un teórico más "puro", de temperamento más sanguíneo, más concienzudo o más curioso, acaso se sienta animado, con nuestro análisis, a proponer una y otra vez

nuevas teorías que compitan, con la esperanza de que una de ellas sea la verídica... aunque jamás podamos estar seguros de que alguna de ellas contenga la verdad.

Así, el teórico puro tiene ante sí más de un camino abierto; y elegirá un método como el de tanteo y eliminación de error sólo si su curiosidad es mayor que su desaliento ante la inevitable incertidumbre e inconclusión de todas nuestras tentativas.

La situación será distinta para el teórico en tanto que hombre de acción. Porque el hombre de acción práctica siempre tiene que *elegir* entre algunas opciones más o menos definidas, ya que *aun la inactividad es una especie de acción*.

Pero toda acción presupone un conjunto de expectativas, es decir, de teorías acerca del mundo. ¿Qué teoría elegirá el hombre de acción? ¿Existe algo que pudiéramos llamar *una elección racional*?

Esto nos lleva a considerar los *problemas pragmáticos de la inducción*, que, para empezar, podemos formular de la siguiente manera:

1) ¿En qué teoría deberíamos basarnos para emprender la acción práctica, desde el punto de vista racional?

2) ¿Qué teoría deberíamos preferir para emprender la acción práctica, desde el punto de vista racional?

Mi respuesta al punto 1) es: desde el punto de vista racional, o racionalista, no deberíamos "basarnos" ni "confiar" en ninguna teoría, pues no se ha demostrado hasta ahora que alguna teoría sea la verídica, o que se pueda demostrar que sea la verídica (o la "confiable").

Mi respuesta al punto 2) es: deberíamos *preferir* la teoría más probada como base para nuestra acción.

En otras palabras, no existe la "confiabilidad absoluta"; pero, puesto que *debemos* elegir, será "racional" elegir la teoría mejor probada. Esto será "lo racional" en el sentido más obvio del mundo para mí: la teoría mejor probada es aquella que, a la luz de nuestra *discusión crítica*, parece la mejor, hasta ese momento; y no conozco nada más "racional" [o racionalista] que una discusión crítica bien hecha.

Como este punto no parece haberse aclarado lo suficiente, trataré de reformularlo aquí de manera algo distinta, que me sugirió David Miller. Olvidémonos por el momento de qué teorías "utilizamos", "elegimos" o "en cuáles basar nuestra acción práctica", y consideremos únicamente la resultante *propuesta o decisión* (hacer X; no hacer X; no hacer nada; etcétera). Tal propuesta, espero, puede criticarse racionalmente; y si somos agentes racionalistas, desearemos que sobreviva, de ser posible, a las más rigurosas pruebas de la crítica que podamos idear. *Pero en tal crítica se utilizarán libremente las teorías científicas mejor probadas de que dispongamos*. Por consiguiente, cualquier propuesta o decisión que

pase por alto estas teorías (huelga decir, cuando estas teorías se aplican al caso) se derrumbará ante la crítica. Si alguna propuesta soporta el peso de la crítica rigurosa, será racional adoptarla.

Esto me parece distar mucho de lo tautológico. Ciertamente, podría objetarse poniendo en tela de juicio la frase que he puesto en cursiva en el párrafo anterior. Podría preguntarse: ¿Por qué utiliza la crítica racionalista las teorías mejor probadas, aunque sean muy poco confiables? Sin embargo, la respuesta sigue siendo la misma. Decidir hacer la crítica de una propuesta práctica desde el punto de vista de la medicina moderna (en vez de hacerla, por ejemplo, en términos frenológicos) es, por sí misma, una decisión *práctica* (en todo caso, tendrá consecuencias prácticas). Así pues, la decisión racionalista es siempre: adoptar métodos de crítica que a su vez se hayan sometido a muy rigurosa crítica, y que hayan sobrevivido.

Por supuesto, en esta formulación existe una regresión infinita; pero es transparentemente inocua.

Ahora bien, no es mi propósito negar (ni afirmar, por cierto) que, al elegir la teoría mejor probada como base para la acción, "confiemos" en ella, en ninguna acepción de esta palabra. Por tanto, acaso pueda describirse como la más "confiable" teoría disponible, en algún sentido de este término. Pero esto no significa que realmente sea "confiable". Es "poco confiable" por lo menos en el sentido de que siempre haremos bien, incluso en la acción práctica, en prever la posibilidad de que algo resulte mal en cuanto a ella misma y en cuanto a nuestras expectativas.

Pero no es sólo esta trivial precaución la que debemos derivar de nuestra respuesta negativa al problema pragmático 1). Es importantísimo, para entender todo el problema, y especialmente lo que he llamado el problema tradicional, que pese a la "racionalidad" de elegir la teoría mejor probada como base de la acción esta elección no sea "racional", en el sentido de que esté basada en *buenas razones en favor* de la expectativa de que en la práctica resultará una elección muy buena: *no puede haber buenas razones en este sentido*, y tal fue precisamente el resultado, la conclusión a la que llegó Hume. Al contrario: aunque sean verídicas nuestras teorías físicas, es muy posible que el mundo tal como lo conocemos, con todas sus regularidades importantes desde el punto de vista pragmático, pueda desintegrarse del todo en el siguiente segundo. Esto debería parecer obvio para todos, hoy día; pero lo dije así⁷ antes de Hiroshima: existe una infinidad de causas probables de desastre local, parcial o total.

⁷ Véase *The Logic of Scientific Discovery* [La lógica de la investigación científica], pp. 253 y ss.

Sin embargo, desde el punto de vista pragmático, la mayoría de estas posibilidades no merecen tomarse en cuenta, *porque no podemos hacer nada* al respecto: están más allá del ámbito de nuestra acción. (Por supuesto, no incluyo a la guerra atómica entre los desastres que están más allá del ámbito de la acción humana, aunque la mayoría de los humanos piensa que esto es así, pues no podemos hacer nada al respecto, como no podemos hacer nada respecto a un acto de Dios.)

Todo esto debe sostenerse, incluso si pudiéramos tener la certeza de que nuestras teorías físicas y biológicas fueran verídicas. Pero no lo sabemos. Al contrario: tenemos buenas razones para sospechar hasta de las mejores de ellas; lo cual, por supuesto, añade más infinitudes a las infinitas posibilidades de que haya una catástrofe.

Son esta clase de consideraciones las que dan tanta importancia a la respuesta negativa de Hume y a la mía propia. Porque ahora podemos ver con toda claridad por qué debemos estar en guardia, so pena de que nuestra teoría del conocimiento pruebe ser demasiado para nosotros. Dicho de manera más precisa, *ninguna teoría del conocimiento debería tratar de explicar por qué tenemos éxito en nuestros intentos de explicar las cosas*.

Incluso si damos por hecho que hemos tenido éxito —que nuestras teorías físicas son verídicas—, podemos aprender de nuestra cosmología cuan infinitamente improbable es tal éxito: nuestras teorías nos dicen que el universo está casi completamente vacío, y que el espacio vacío está lleno de radiación caótica. Y casi todos los lugares que no están vacíos están ocupados, ya sea por polvo caótico, o por gases, o por estrellas muy calientes, todo ello en condiciones que parecen hacer imposible la aplicación de algún método físico para adquirir conocimiento.

Existen muchos mundos, posibles y reales, en los que fallaría la búsqueda de conocimiento y de regularidades. E incluso en el mundo tal como lo conocemos en la realidad a través de las ciencias, el que haya condiciones bajo las cuales la vida y la búsqueda de conocimiento pudieran darse —y tener éxito— parece ser algo casi infinitamente improbable. Además, al parecer, si se dieran algún día esas condiciones, estarían sujetas a desaparecer otra vez, después de un tiempo que, hablando en términos cosmológicos, es muy breve.

En este sentido, la inducción es inductivamente inválida, como ya lo he dicho antes. Esto significa que cualquier respuesta positiva al problema lógico de Hume (por ejemplo, la tesis de que la inducción sea válida) resultaría paradójica. Porque, por una parte, si la inducción es el método de la ciencia, entonces la cosmología moderna está, por lo menos aproximadamente, en lo cierto (no disputo esto); y por otra parte, la eos-

mología moderna nos enseña que generalizar a partir de observaciones tomadas en su mayoría en nuestra región increíblemente idiosincrásica del universo sería casi siempre bastante inválido. Así pues, si la inducción es "inductivamente válida", casi siempre nos llevará a obtener conclusiones falsas; y, por tanto, es inductivamente inválida.

8. EL PROBLEMA DE LA DEMARCACIÓN (1974)

I. LA CIENCIA CONTRA LA NO CIENCIA

ME OCUPARÉ ahora del *problema de la demarcación*, y de explicar cómo se relaciona este problema con los problemas del contenido empírico y de la comprobabilidad.

Los grandes científicos, como Galileo, Kepler, Newton, Einstein y Bohr (por sólo citar a los ya fallecidos) representan para mí una idea simple, pero impresionante, de la ciencia. Es obvio que una lista como ésta, por muy larga que fuese, no podría *definir al científico* ni a la ciencia, *in extenso*. Pero me sugiere una sobresimplificación, de la que, en mi opinión, podemos aprender mucho. Es la obra de los grandes científicos la que tengo como paradigma de la ciencia. No es que no respete también a los científicos de menor talla; hay cientos de grandes hombres y de grandes científicos que entran en la categoría casi heroica.

Pero, dicho sea con todo respeto para los científicos de menor talla, deseo transmitir aquí una idea heroica y romántica de la ciencia y de sus trabajadores: hombres que se dedican humilde y devotamente a buscar la verdad, a hacer que crezca nuestro conocimiento; hombres cuya vida consistió en una aventura de ideas osadas. Estoy dispuesto a considerar entre ellos a muchos de sus menos brillantes ayudantes que se dedicaron igualmente a buscar la verdad, la gran verdad. Pero no cuento entre ellos a aquellos para quienes la ciencia no es más que una profesión como cualquier otra, una técnica: a aquellos que no están motivados profundamente por grandes problemas y por las sobresimplificaciones de las soluciones osadas.

Es la ciencia en este sentido heroico la que deseo estudiar. Como resultado colateral, descubro que podemos arrojar mucha luz incluso en los más modestos trabajadores de las ciencias aplicadas.

Entonces, esto, para mí, es la ciencia. No intento definirla, y por muy buenas razones. Sólo deseo esbozar un cuadro sencillo de la clase de hombres en los que pienso cuando hablo de la ciencia, y de sus actividades. Y el cuadro será una sobresimplificación: son hombres de osadas ideas, pero críticos muy rigurosos de sus propias ideas; tratan de averiguar si sus ideas están en lo correcto, y antes que nada intentan averiguar si no serán, acaso, erróneas. Trabajan con osadas conjeturas y con severos intentos de refutar sus propias conjeturas.

Mi criterio de demarcación entre la ciencia y la no ciencia es un sencillo análisis lógico de este cuadro. Cuan bueno o cuan malo sea este cuadro, lo demostrará su fertilidad.

Las ideas audaces son nuevas, intrépidas, hipótesis o conjeturas. Y también los intentos severos de refutarlas representan discusiones críticas rigurosas y severas pruebas empíricas.

¿Cuándo es una conjetura osada, y cuándo no lo es, en el sentido que aquí propongo? Respuesta: es osada, siempre y cuando corra el gran riesgo de resultar falsa: si las cosas podrían ser de otra manera, y al mismo tiempo parecen ser de otra manera.

Consideremos un sencillo ejemplo. La conjetura de Copérnico o de Aristarco de que el Sol, y no la Tierra, está en el centro del universo, era una idea increíblemente osada. De paso, digamos que era una idea falsa; nadie acepta hoy la conjetura de que el Sol está (en el sentido que daban a esta idea Aristarco y Copérnico) inmóvil en el centro del universo. Pero esto no afecta a la osadía de la conjetura, ni a su fertilidad. Y una de sus principales consecuencias —que la Tierra no está inmóvil en el centro del universo, sino que tiene (por lo menos) un movimiento diario y otro anual— todavía se acepta plenamente, pese a algunas malas interpretaciones de la relatividad. •

Pero no es la actual aceptación de la teoría lo que deseo analizar, sino su osadía. Era osada, porque chocaba con todas las opiniones que entonces se aceptaban, y porque contradecía la evidencia *prima facie* de los sentidos corporales. Era osada, porque postulaba una hasta entonces oculta realidad detrás de las apariencias.

No era osada en otro sentido muy importante: ni Aristarco ni Copérnico sugirieron un experimento factible y crucial. De hecho, no sugirieron que nada estuviera mal en cuanto a las apariencias tradicionales: dejaron de lado, sin más, las apariencias aceptadas: sólo las reinterpretaron. No ansiaban arriesgar el cuello al predecir nuevas apariencias observables. (Esto constituye una sobresimplificación respecto de Copérnico, pero es casi una certeza respecto de Aristarco.)

En la medida en que esto sea cierto, las teorías de Aristarco y de Copérnico pueden describirse con mi terminología como no científicas, o metafísicas. En la medida en que Copérnico sí hizo un pequeño número de predicciones no muy importantes, su teoría es, según mi terminología, científica. Pero incluso como teoría metafísica, distaba mucho de no tener significación; y al proponer una nueva y audaz visión del uni-

verso, hizo una gran contribución para el advenimiento de la nueva ciencia.

Kepler avanzó mucho más. Él también tenía una visión metafísica, en parte basada en la teoría copernicana, de la realidad del mundo. Pero esta visión suya lo llevó a hacer muchas predicciones detalladas acerca de las apariencias. Al principio, estas predicciones no concordaban con las observaciones. Trató de reinterpretar las observaciones a la luz de sus teorías; pero su adicción a la búsqueda de la verdad era mayor que su entusiasmo por la armonía metafísica del mundo. Así pues, se vio obligado a renunciar a varias de sus teorías favoritas, una tras otra, para sustituirlas con otras que se apegaban a los hechos. Fue una lucha encarnizada. El resultado final, sus famosas e importantísimas tres leyes, no fueron de su agrado, excepto la tercera. Mas soportaron las severas pruebas a que las sometió, pues concordaban con las detalladas apariencias, las observaciones que él había heredado de Tycho Brahe.

Las leyes de Kepler constituyen excelentes aproximaciones a lo que hoy pensamos que son los verdaderos movimientos de los planetas de nuestro sistema solar. Constituyen, incluso, excelentes aproximaciones a los movimientos de distantes sistemas de estrellas binarias que desde entonces se han descubierto. No obstante, son meras *aproximaciones* a lo que parece ser la verdad; *no son verídicas*.

Se han probado a la luz de nuevas teorías —las de Newton y las de Einstein— que predijeron pequeñas desviaciones de las leyes de Kepler. (Según Newton, las leyes de Kepler son válidas sólo para los sistemas siderales binarios [véase el escrito 12, más adelante].) Así, los cruciales experimentos resultaron contrarios a Kepler, muy ligeramente, pero con suficiente claridad.

De estas tres teorías —la de Kepler, la de Newton y la de Einstein— la última y de mayor éxito aún es la de Einstein; y fue esta teoría la que me orientó hacia la filosofía de la ciencia. Lo que tanto me impresionó acerca de la teoría de la gravitación de Einstein fueron los siguientes puntos:

1) Era una teoría muy osada. Se desviaba mucho en su concepto fundamental de la teoría de Newton, que en aquel tiempo tuvo muchísimo éxito. (Las pequeñas desviaciones del perihelio de Mercurio no molestaron mucho a nadie, a la luz de su otro casi increíble éxito. Que hubieran causado mucha discrepancia estas desviaciones, ya es otro asunto.)

2) Desde el punto de vista de la teoría de Einstein, la teoría de Newton era una excelente aproximación, aunque falsa (así como desde el punto de vista de la teoría de Newton, la de Kepler y la de Galileo eran excelentes aproximaciones, aunque falsas). Por lo visto, no es su veracidad lo que decide acerca de la índole científica de una teoría.

¹ Véase *Conjectures and Refutations*, p. 110: "Desde el punto de vista de la relatividad general, [...] la Tierra gira [...] precisamente en el sentido en que gira una rueda de bicicleta".

3) Einstein derivó de su teoría tres importantes predicciones de efectos muy observables y diferentes, en dos de los cuales nadie había pensado antes que él, y que contradecían, todas, la teoría de Newton, en la medida en que podía afirmarse que entraban en el campo de aplicación de la teoría de Einstein.

Pero quizá lo que más me haya impresionado hayan sido los dos puntos siguientes:

4) Einstein declaró que estas predicciones eran de importancia capital: si no concordaban con sus precisos cálculos teóricos, consideraría refutada su teoría.

5) Pero, aunque se observaron según había predicho, Einstein declaró que *su teoría era falsa*: declaró que sería una mejor aproximación a la verdad que la teoría de Newton, y dio sus razones por las que no consideraría verídica su nueva teoría, incluso si todas las predicciones se realizaran. Esbozó un número de exigencias que tendría que satisfacer una verdadera teoría (del campo unificado) y declaró que su teoría era cuando mucho una aproximación a esta todavía no lograda teoría del campo unificado.

Podemos observar, de paso, que Einstein, al igual que Kepler, fracasó en su sueño científico... o en su sueño metafísico; en este contexto, no importa qué etiqueta utilicemos. Lo que hoy llamamos las leyes de Kepler o la teoría de la gravitación de Einstein son resultados que de ningún modo satisficieron a sus creadores, que siguieron trabajando cada cual en su sueño hasta el final de sus existencias. Y podemos señalar un punto similar a éste en Newton: nunca creyó que una teoría de acción a distancia pudiera ser a la postre una aceptable explicación de la gravedad.²

La teoría de Einstein fue probada por primera vez con el famoso experimento de Eddington, del eclipse de 1919. Pese a que no creía en la verdad de su teoría, de que sólo era una nueva e importante aproximación a la verdad, Einstein nunca dudó del resultado de este experimento: la coherencia interior, la lógica interna de su teoría, lo convencieron de que se trataba de un paso hacia adelante, aunque pensara que no podía ser verídica. Desde entonces, ha pasado por una serie de pruebas posteriores, todas con gran éxito. Pero algunas personas piensan todavía que la concordancia entre la teoría de Einstein con las observaciones puede ser resultado de (increíblemente improbables) accidentes. Es imposible que nos pronunciemos en esto a favor o en contra; sin embargo, podemos decir que tal concordancia podría ser resultado de que la teoría de Einstein a este respecto fuera una aproximación fantásticamente buena a la verdad.³

² Véase *Conjectures and Refutations*, capítulo 3, notas 20-22, pp. 106 y ss.

³ Véase también mi ensayo "The Present Significance of Two Arguments of Henri Poincaré" ["La significación actual de dos argumentos de Henri Poincaré"], en *Methodology and Science* 14, 1981, pp. 260-264.

El cuadro de la ciencia del que sólo he mencionado algunos rasgos puede trazarse como sigue.

Hay una realidad detrás del mundo tal como aparece ante nosotros, o posiblemente una realidad compuesta de muchas capas, y de la que las apariencias son las capas más externas. Lo que hace el gran científico es conjeturar osadamente, con atrevidas suposiciones, cómo son estas realidades interiores. Lo cual es afín a la creación de mitos. (Históricamente, podemos seguir hacia atrás en el tiempo las ideas de Newton, pasando por Anaximandro, hasta Hesíodo, y las ideas de Einstein, a través de Faraday, Boscovich, Leibniz y Descartes, hasta Aristóteles y Parménides.)⁴ La osadía de estas ideas puede medirse mediante la distancia entre el mundo de la apariencia y la realidad conjeturada, las hipótesis explicativas.

Pero existe otra osadía, de clase especial: *la osadía de predecir* aspectos del mundo de la apariencia que hasta ese momento se han pasado por alto, pero que debe tener el mundo, si la realidad conjeturada es (más o menos) la correcta, si las hipótesis explicativas son (aproximadamente) verídicas. Pienso en esta osadía más especial cuando hablo de conjeturas científicas osadas. Es la osadía de una conjetura que corre un riesgo real: el riesgo de que se le ponga a prueba, y de que se le refute; el riesgo de que choque con la realidad.

Así pues, mi propuesta era, y es, que es esta segunda osadía, junto con la disposición a buscar pruebas y refutaciones, la que distingue a la ciencia "empírica" de la no ciencia, y especialmente, de los mitos precientíficos y metafísicos.

Llamaré a esta propuesta D): D) que representa "demarcación".

La propuesta D), que he puesto en caracteres cursivos, es lo que todavía considero el centro de mi filosofía. Pero siempre he sido muy crítico de cualquier idea mía; por ello, he tratado de encontrar inmediatamente fallas a esta idea particular, años antes de publicarla. Y la publiqué junto con los resultados de esta crítica. Mi crítica me llevó a una secuencia de mejoras o refinamientos de la propuesta D): no fueron concesiones posteriores, sino que se publicaron junto con la propuesta, como parte integral de ella.⁵

⁴ Véase el índice de *Conjeturas y refutaciones*, bajo estos nombres, y *Teoría cuántica y el cisma en física*, sección 20.

⁵ Debo hacer hincapié en este punto, porque Ayer ha aseverado en las pp. 583 y ss de "Philosophy and Scientific Method" ["Filosofía y el método científico"], en *Proceedings of the XIVth International Congress of Philosophy*, Viena, del 2 al 9 de septiembre de 1968, volumen 1, pp. 536-542, que "En los tiempos modernos dos tesis han dominado en este campo. Según la primera, lo que se necesita es que la hipótesis sea verificable; según la otra, que sea falsificable". Y después de esbozar muy brevemente una historia del criterio de verificabilidad, escribe: "En su forma actual, lo único que se exige de una hipótesis

II. DIFICULTADES RELATIVAS A LA PROPUESTA DE DEMARCACIÓN

1) Desde un principio llamé a mi criterio de demarcación una *propuesta*. Esto se debió, en parte, a la incomodidad y el desagrado que me inspiran las definiciones. Las definiciones son, o abreviaciones, y por tanto, innecesarias, aunque quizá, convenientes, o bien intentos aristotélicos de "expresar la esencia" de una palabra y, por consiguiente, dogmas inconscientes y convencionales [véase, más arriba, el escrito 6]. Si definimos el vocablo "ciencia" según mi criterio de demarcación (reconozco que esto es más o menos lo que estoy haciendo), entonces cualquiera podría proponer otra definición, como: "la ciencia es la suma total de aseveraciones verídicas". La discusión de los méritos de tales definiciones puede resultar inútil. Por ello he hecho aquí, al principio, una descripción de la gran ciencia, o de la ciencia heroica, y luego propuse un criterio que nos permite demarcar —en términos muy generales— esta clase de ciencia. Cualquier demarcación en el sentido de mi propuesta debe ser muy *general*. (Ésta es una de las grandes diferencias que la apartan de cualquier criterio formal de significado, de cualquier "lenguaje de la ciencia" artificial.) Porque la transición entre la metafísica y la ciencia no es tajante: lo que ayer fue una idea metafísica, puede convertirse mañana en una teoría científica comprobable; y esto sucede con frecuencia (di varios ejemplos de esto en *La lógica de la investigación científica* y en otras obras mías: el atomismo acaso sea el mejor ejemplo).

Así pues, una de las dificultades consiste en que nuestro criterio no debe ser demasiado estricto; y en el capítulo "Grados de comprobabilidad" de mi obra *La lógica de la investigación científica* sugerí (como una especie de segunda mejora del criterio *D*) de la sección anterior) que una teoría es científica según su grado de comprobabilidad.

Esto, incidentalmente, redundó después en uno de los más fructíferos

científica es que figure de manera no trivial en una teoría que esté abierta a la confirmación cuando se considere como un todo".

"En el caso del principio de falsificabilidad —prosigue Ayer—, el proceso de adaptación ha sido menos explícito. Algunos de sus defensores hablan todavía como si la formulación que le ha dado el profesor Popper en los primeros capítulos de su *La lógica de la investigación científica* siguiera siendo válida. Sin embargo, el hecho es que el profesor Popper mismo ha creído necesario modificarla en la revisión de este mismo libro." A esto sólo puedo replicar que: 1) me parece mejor introducir las necesarias modificaciones "en este mismo libro" en que se hizo la proposición; 2) introduje la falsificabilidad como un criterio de demarcación, en la p. 40 de *The Logic of Scientific Discovery* [*La lógica de la investigación científica*] y "creí necesario" esbozar todas las diferentes objeciones en la página siguiente, en la misma sección, al anunciar mi intención de analizar con mayor detalle cada una de ellas posteriormente; 3) la única dificultad que pospuse —la no falsificabilidad formal de las afirmaciones de probabilidad— se resolvió mediante una propuesta metodológica.

descubrimientos de ese libro: que hay grados de comprobabilidad (o de carácter científico) que pueden identificarse con grados de contenido empírico (o de contenido informativo).

2) La fórmula *D*) de la anterior sección está expresada en lenguaje un tanto psicológico. Podemos mejorarla mucho si hablamos de *sistemas teóricos* o de *sistemas de declaraciones o de aseveraciones*, como lo hice en todo el texto de *La lógica de la investigación científica*. Esto nos lleva a reconocer inmediatamente uno de los problemas relacionados con el criterio de falsificabilidad de la demarcación: aunque podamos aplicarlo a *sistemas* de aseveraciones o declaraciones, puede ser difícil, si no imposible, dilucidar qué aseveración particular, o qué subsistema de un sistema de aseveraciones, se ha expuesto a determinada prueba experimental. Así, podemos describir un *sistema* como científico, o empíricamente comprobable, aunque al mismo tiempo tengamos incertidumbre acerca de sus partes constituyentes.

Un ejemplo de lo anterior es la teoría de la gravitación de Newton. Se ha preguntado a menudo si las leyes del movimiento de Newton, o cuáles de ellas, serán definiciones enmascaradas, en vez de aseveraciones empíricas.

Mi respuesta a esta pregunta es la siguiente: La teoría de Newton es un sistema. *Si la declaramos falsa, si comprobamos que es falsa, tendremos que falsificar todo el sistema*. Acaso le echemos la culpa de esta falsedad a una o a otra de sus leyes. Pero esto sólo significa que *conjeturamos* que cierto cambio en el sistema lo liberará de su calidad de falso; o, en otras palabras, que conjeturamos que cierto sistema alterno será una mejora, una mejor aproximación a la verdad.

Pero también hay que tener en cuenta esto: atribuir la culpa de la comprobación de falsedad a cierto subsistema es una típica hipótesis, una conjetura como cualquier otra, o quizá apenas algo más que una sospecha vaga, si no se ha hecho una sugerencia de sistema alterno. Y lo mismo se aplica a lo contrario: la decisión de que cierto subsistema no ha de considerarse culpable de la falsedad comprobada es también una típica conjetura. La atribución o la no atribución de responsabilidad por el fracaso es conjetural, como lo es todo en ciencia; y lo que importa es la propuesta de un sistema conjetural alterno, competitivo, que soporte la prueba de detección de falsedad.

3) Los puntos 1) y 2) ilustran que, por muy correcto que sea mi criterio de las conjeturas osadas y rigurosas refutaciones, hay dificultades que no debemos pasar por alto. Una dificultad primitiva de este tipo puede describirse como sigue: Un biólogo presenta la conjetura de que todos los cisnes son blancos. Cuando se descubren cisnes negros en Australia, él dice que no lo han refutado; insiste en que esos cisnes negros

son una especie nueva de aves, ya que es *parte de la propiedad definitoria* del cisne que debe ser blanco. En otras palabras, puede escapar de la refutación, aunque pienso que posiblemente aprendería más si reconociera haberse equivocado.

En cualquier caso —y esto es muy importante—, la teoría de que "Todos los cisnes son blancos" es refutable por lo menos en este sentido lógico y claro: debe declararla refutada cualquiera que acepte que hay por lo menos un cisne que no es blanco.

4) El principio implícito en este ejemplo es muy primitivo, pero tiene gran cantidad de aplicaciones. Los científicos se han inclinado desde hace mucho por considerar los pesos atómicos de los elementos, los puntos de fusión y otras semejantes propiedades como *definitorias* de los materiales: no puede haber ninguna clase de agua cuyo punto de congelamiento [al nivel del mar] no sea de 0°C; no sería agua, por más que en algunos aspectos podría ser semejante al agua. Pero, si es así, según mi criterio de demarcación, "El agua se congela al nivel del mar a 0°C" no sería una aseveración científica o empírica; sería una tautología: parte de una definición.

Está claro que hay aquí un problema: o mi criterio de demarcación es refutado, o hemos de aceptar la posibilidad de descubrir cierta agua cuyo punto de congelación no sea 0°C al nivel del mar.

5) Por supuesto, definiendo la segunda posibilidad, y sostengo que a partir de este sencillo ejemplo podemos aprender mucho sobre las ventajas de mi propuesta D). Porque, supongamos que hemos descubierto cierta agua con distinto punto de congelación. ¿Podríamos seguir llamando a esta agua, "agua"? *Afirmo que la pregunta es del todo impropcedente*. La hipótesis científica era que un líquido (sin importar cómo lo llamemos) con una considerable lista de propiedades químicas y físicas se congela a 0°C. Si cualquiera de estas propiedades que se han conjeturado como constantemente presentes en ese líquido no se materializara, entonces debemos aceptar que *estábamos equivocados*; y así *se abren nuevos e interesantes problemas*. El menor de ellos sería si podríamos seguir llamando "agua" a ese líquido: *esto* es puramente arbitrario o convencional. Así, mi criterio de demarcación no sólo no ha sido refutado con este ejemplo: nos ha ayudado a descubrir qué es importante para la ciencia y qué es arbitrario e impropcedente.

6) Como ya expliqué en el primer capítulo de *La lógica de la investigación científica*, podemos siempre adoptar tácticas evasivas ante las refutaciones. Por razones históricas, originalmente llamé a estas tácticas "estratagemas convencionalistas" [o rodeos, "desviaciones"]; pero ahora las llamo "*tácticas o estratagemas de inmunización*":⁶ siempre

podemos inmunizar una teoría para protegerla de la refutación. Existen muchas tácticas inmunizadoras de este tipo; y si no se nos ocurre algo mejor, podemos siempre negar la objetividad —o incluso la existencia— de las observaciones refutadoras. (Recordemos a la gente *que se negaba* a mirar por el telescopio de Galileo.) Los intelectuales a quienes interesa más estar en lo cierto que aprender algo interesante, pero inesperado, no son, por cierto, raras excepciones.

7) Ninguna de las dificultades que hasta ahora hemos analizado es terriblemente grave: al parecer, sólo un poco de honradez intelectual las superaría. Esto es verdad, en términos generales. Pero, ¿cómo describir esta probidad intelectual en términos de lógica? La describí en *La lógica de la investigación científica* como una *regla de método*, o *regla metodológica*: "No trates de evadir la falsificación, ¡pero procura no caer en motivo de que declaren falsa tu teoría!"

8) Pero fui un poco más autocrítico: primero noté que tal regla o método es, necesariamente, algo vago..., como también lo es el problema de la demarcación. Claro está que podemos decir que si evitamos la falsificación *a cualquier precio*, renunciamos a la ciencia empírica, en la acepción que yo le doy. Pero también descubrí que la hipersensibilidad respecto a refutar la crítica era igualmente peligrosa: hay un lugar legítimo para el dogmatismo, aunque un lugar muy reducido. Quien renuncia demasiado fácilmente a su teoría frente a aparentes refutaciones, jamás descubrirá las posibilidades inherentes a su teoría. *En ciencia, siempre hay lugar para el debate*: para el ataque y, por ende, para la defensa. Sólo si intentamos defenderlas, aprenderemos todas las diferentes posibilidades inherentes a nuestras teorías. Como siempre, la ciencia es conjetura. Tenemos que conjeturar cuándo dejar de defender una teoría favorita, y cuándo intentar la elaboración de otra que la sustituya.

9) Por tanto, no propuse la regla simple: "Busque usted refutaciones, y nunca defienda dogmáticamente su teoría." Pero esta regla habría sido mucho mejor consejo que aconsejar la defensa dogmática a toda costa. La verdad es que debemos ejercer constantemente la crítica; debemos ser autocríticos respecto a nuestras propias teorías, y autocríticos respecto a nuestras propias críticas; y, por supuesto, jamás debemos evitar un debate.

Así pues, tal es en términos generales *la forma metodológica de D)*, del criterio de demarcación. Proponga usted teorías que puedan criticarse. Piense en posibles experimentos que se puedan falsificar; en experimentos cruciales. Pero no renuncie usted demasiado fácilmente a sus teorías, en todo caso, no antes de haber examinado con espíritu crítico sus propias críticas.

⁶ Este término se debe a Hans Albert.

III. TEORÍAS EMPÍRICO-CIENTÍFICAS Y TEORÍAS NO CIENTÍFICAS

Las dificultades relacionadas con mi criterio de demarcación *D*) son importantes, pero no hay que exagerarlas. Es un criterio vago, ya que es un criterio metodológico, y *porque la demarcación entre la ciencia y la no ciencia es vaga*. Pero es lo suficientemente certero para establecer una distinción entre muchas teorías de física, por una parte y, por la otra, las teorías metafísicas, tales como el psicoanálisis o el marxismo (en su forma actual). Ésta es, por supuesto, una de mis tesis principales; y nadie que no la haya comprendido puede decir que haya entendido mi teoría.

La situación respecto del marxismo es, incidentalmente, muy distinta de la situación con el psicoanálisis; el marxismo fue en otro tiempo una teoría científica: predijo que el capitalismo llevaría a la creciente miseria y, mediante una revolución más o menos moderada, al socialismo; predijo que esto sucedería primero en los países más técnicamente desarrollados; y predijo que la evolución tecnológica de los "medios de producción" redundaría en los desarrollos sociales, políticos e ideológicos, y no al revés.

Pero la (así llamada) revolución socialista ocurrió primero en uno de los países más atrasados en lo tecnológico. Y en vez de que los medios de producción crearan una nueva ideología, fue la ideología de Lenin y de Stalin la que preconizó que Rusia debía seguir adelante con la industrialización ("El socialismo es la dictadura del proletariado, más la electrificación"), que promovió el nuevo desarrollo de los medios de producción.

Así pues, podríamos decir que el marxismo fue alguna vez una ciencia, pero una ciencia a la que refutaron algunos de los hechos que contradijeron sus predicciones (aquí me he referido sólo a unos cuantos de estos hechos).⁷

Sin embargo, el marxismo ya no es una ciencia; porque infringió la regla de que debemos aceptar la falsificación y se inmunizó a sí mismo contra las más flagrantes refutaciones de sus predicciones. Desde ese momento, por tanto, puede describirse sólo como una no ciencia..., como un sueño metafísico, si se quiere, maridado a una cruel realidad.

El caso del psicoanálisis es muy diferente. Constituye una interesante metafísica psicológica (y no cabe duda de que hay alguna verdad en él, como sucede tan a menudo en las ideas metafísicas), pero nunca fue una ciencia. Puede haber muchísimas personas que sean casos freudianos o

⁷ Para una discusión más detallada, véase *The Open Society and Its Enemies [La sociedad abierta y sus enemigos]*, vol. II, pp. 108 y ss [y también el escrito 26 de la presente obra].

adlerianos: Freud mismo era claramente un caso freudiano, y Adler, un caso adleriano. Pero lo que les impide a sus teorías ser científicas en el sentido que he descrito aquí es, muy sencillamente, que no excluyen ninguna conducta humana físicamente posible. Cualquier cosa que alguien haga es en principio explicable en términos freudianos o adlerianos. (El rompimiento de Adler con Freud fue más adleriano que freudiano, pero Freud jamás lo consideró una refutación de su propia teoría.)

El punto está muy claro: ni Freud ni Adler excluyen a ninguna persona en particular actuando de ningún modo particular, cualesquiera que sean las circunstancias exteriores. Que un hombre sacrificara la vida para salvar a un niño que estuviera ahogándose (un caso de sublimación), o que asesinara al niño ahogándolo (un caso de represión), no podría predecirse o excluirse con la teoría de Freud; *la teoría freudiana era compatible con cualquier cosa que hubiera ocurrido..., hasta sin ningún tratamiento inmunizador especial*.

Así pues, mientras que el marxismo se volvió no científico al adoptar una estrategia inmunizante, el psicoanálisis, desde un principio, era ya inmune, y siguió siéndolo.⁸ En contraste con esto, la mayoría de las teorías de física están bastante libres de tácticas inmunizadoras, y, para empezar, *son muy falsificables*. En general, *excluyen una infinidad de concebibles posibilidades*.

El valor principal de mi criterio de demarcación era, por supuesto, señalar estas diferencias. Y me llevó a elaborar la teoría de que el contenido empírico de una teoría podría medirse por el número de posibilidades que excluyera (siempre y cuando se adoptara una metodología no inmunizante en medida razonable).

IV. HIPÓTESIS *AD HOC* E HIPÓTESIS AUXILIARES

Hay un método importante para evitar o soslayar las refutaciones: es el método de las hipótesis *ad hoc* y de las hipótesis auxiliares.

Si alguna de nuestras conjeturas resulta falsa —por ejemplo, si el planeta Urano no se desplaza exactamente como lo exige la teoría de Newton— *entonces tenemos que cambiarla teoría*. Pero, principalmente, puede haber dos tipos de cambios en este sentido: *los cambios conservadores y los cambios revolucionarios*. Y entre los cambios más conservadores puede haber otros dos: *las hipótesis ad hoc y las hipótesis auxiliares*.

En el caso de las desviaciones en el desplazamiento de Urano, la hipótesis adoptada fue parcialmente revolucionaria: lo que se conjeturó

⁸ Véase *Conjectures and Refutations*, capítulo 1, especialmente las pp. 35-38.

fue la existencia de un nuevo planeta, lo cual no afectó a las leyes de los movimientos planetarios de Newton, pero que sí afectó al mucho más antiguo "sistema del mundo". La nueva conjetura fue más auxiliar que *ad hoc*, pues aunque sólo había esta razón *ad hoc* para proponerla, era *independientemente estable*: la posición del nuevo planeta (Neptuno) se calculó, el planeta hasta entonces desconocido se descubrió ópticamente, y se vio que explicaba muy bien las anomalías de Urano. Así, la hipótesis auxiliar siguió dentro del marco teórico newtoniano, y la amenazante refutación se convirtió en un resonante éxito.

Llamo conjetura *ad hoc* si se propone (como ésta que he ejemplificado) para explicar determinada dificultad, pero que (en contraste con ésta) *no pueda comprobarse independientemente*.

Resulta claro que, como todo, en metodología, la distinción entre una hipótesis *ad hoc* y una hipótesis conservadora auxiliar es un tanto vaga. Pauli propuso la hipótesis del neutrino bastante conscientemente como una hipótesis *ad hoc*. Al principio, no tenía ninguna esperanza de que algún día se descubriera una prueba independiente de la existencia del neutrino; en aquel tiempo, esto parecía imposible. Así que aquí tenemos el ejemplo de una hipótesis *ad hoc* que, con el desarrollo del conocimiento, perdió su carácter de *ad hoc*. Y esto nos advierte que no debemos pronunciar un edicto demasiado severo contra las hipótesis *ad hoc*: pueden convertirse en hipótesis comprobables, como también puede ocurrir a las hipótesis metafísicas. Pero, en general, nuestro criterio de comprobabilidad nos pone sobre aviso en contra de las hipótesis *ad hoc*; y Pauli mismo, al principio, distaba mucho de estar satisfecho con su conjetura del neutrino, que se habría abandonado, con toda seguridad, si a la postre unos nuevos métodos no hubieran permitido hacer pruebas independientes de su existencia.

Las hipótesis *ad hoc* —es decir, en su tiempo hipótesis auxiliares no comprobables— pueden salvar de una refutación *particular* a casi cualquier teoría. Pero esto no significa que podamos seguir proponiendo hipótesis *ad hoc* cuanto lo deseemos. Puede convertirse en hipótesis comprobable; y una prueba que resulte negativa acaso nos obligue, o a renunciar a ella, o a proponer una nueva hipótesis *ad hoc* secundaria, y así sucesivamente, *ad infinitum*. En realidad, esto es algo que debemos evitar casi siempre. (Digo "casi siempre", porque las reglas metodológicas no son inflexibles ni de rápida aplicación.)

Además, no debemos exagerar la posibilidad de arreglar las cosas mediante hipótesis *ad hoc*: hay muchas refutaciones que no pueden soslayarse de esta manera, aunque puede darse siempre algún tipo de táctica inmunizante, como hacer caso omiso de la refutación.