

SIMÓN BLACKBURN

PENSAR

Una incitación a la filosofía

PAIDÓS

Barcelona • Buenos Aires • México

Título original: *Think. A compelling introduction to philosophy*

Originalmente publicado en inglés, en 1999, por Oxford University Press, Oxford, R.U.

Traducción publicada de conformidad con Oxford University Press

Originally published in English in 1999 is published by arrangement with Oxford University Press

Traducción de Ramón Vilà Vernis

Cubierta de Mario Eskenazi

© 1999 Simón Blackburn

© 2001 de la traducción, Ramón Vilà

© 2001 de todas las ediciones en castellano

Ediciones Paidós Ibérica, S. A.,

Mariano Cubí, 92 - 08021 Barcelona

y Editorial Paidós, SAICF,

Defensa, 599 - Buenos Aires.

<http://www.paidos.com>

ISBN: 84-493-1037-7

Depósito legal: B-10.166/2001

Impreso en A&M Gràfic, S.L.

08130 - Santa Perpetua de Mogoda (Barcelona)

Impreso en España - Printed in Spain

Sumario

Prefacio.....	4
Introducción.....	6
1.El conocimiento.....	18
2.La mente.....	51
3.La libertad.....	81
4.El yo.....	119
5.Dios.....	147
6.Razonar.....	189
7.El mundo.....	229
8.¿Qué hacer?.....	265
Notas.....	293
Bibliografía.....	299

Prefacio

Este libro se ha gestado a lo largo de años de lidiar con el problema de conseguir que la gente se interese por las ideas. Me he entregado a esta tarea ante todo como profesor, aunque también me he esforzado por explicar el valor de las humanidades en general, y de la filosofía en particular, a una audiencia más amplia. Sin duda, estoy en deuda en primer lugar con el clima intelectual de nuestra época, cuyo escepticismo hacia el valor de la educación superior hizo evidente para mí lo urgente de la tarea. Una segunda deuda, más seria, se refiere a mis alumnos de todos estos años, cuyos gestos de asentimiento o incompreensión han dado al libro su forma definitiva. También debo agradecer la experiencia de primera mano que me brindaron algunos profesores asistentes que trabajaron con sus alumnos sobre versiones anteriores del libro, aquí en la Universidad de Carolina del Norte. Sin embargo, nunca me hubiera atrevido a dar el paso definitivo si no fuera por el generoso apoyo que me prestaron Catherine Clarke y Angus Phillips, de Oxford University Press. Angus ha seguido de cerca el progreso de mi trabajo y debo mucho a su aliento y consejo.

Huw Price y Ralph Walker, cuyas sugerencias fueron de inestimable valor, leyeron versiones anteriores del texto. Yuri Balashov y Dan Ryder me prestaron su ayuda en materias específicas. En beneficio de la brevedad no he incluido un glosario de términos filosóficos, que en todo caso no habría hecho más que recordar las definiciones que figuran en mi *Oxford Dictionary of*

Philosophy.

El magnífico trabajo de edición de Maura High y Angela Blackburn me hizo tomar conciencia, para mayor incomodidad mía, de mis propias limitaciones como escritor, al tiempo que afortunadamente las ocultaba al público en general. Por supuesto, Angela tuvo que soportar los inconvenientes habituales de tener un marido escritor, y sin su apoyo nada habría sido posible.

Introducción

Este libro está destinado a la gente que desea reflexionar sobre los grandes temas: el conocimiento, la razón, la mente, la libertad, el destino, la identidad, Dios, la bondad, la justicia. Tales cuestiones no pertenecen al coto privado de los especialistas. Son cuestiones sobre las que hombres y mujeres se preguntan naturalmente, ya que estructuran nuestra forma de pensar acerca del mundo y acerca del lugar que ocupamos en él. Son también temas sobre los cuales los pensadores han tenido cosas que decir. En este libro pretendo mostrar diversas formas de pensar sobre los grandes temas, así como algunas de las cosas que los pensadores han dicho sobre ellos. Los lectores que hayan asimilado el libro deberían encontrarse en mejor posición para abordar estas cuestiones. También deberían ser capaces de leer con placer y un grado razonable de comprensión a grandes autores que de otro modo podrían resultar desconcertantes.

La palabra «filosofía» trae consigo connotaciones desafortunadas: improductiva, etérea, rebuscada. Sospecho que todos los filósofos y estudiantes de filosofía compartimos aquel momento de embarazoso silencio que se produce cuando alguien nos pregunta inocentemente qué es lo que hacemos. Yo prefiero presentarme como un ingeniero de conceptos. El filósofo estudia la estructura del pensamiento del mismo modo en que el ingeniero estudia la estructura de los objetos materiales. Comprender una estructura significa identificar cómo funcionan sus partes y cómo se relacionan entre sí. También implica saber qué sucedería, para mejor o para peor, en caso de que se

introdujeran cambios. Éste es también nuestro objetivo cuando investigamos las estructuras que configuran nuestra visión del mundo. Nuestros conceptos o ideas constituyen el edificio mental en el que vivimos. Puede que nos sintamos orgullosos de las estructuras que hemos construido, o bien podemos convencernos de que debemos desmantelarlas y empezar otra vez desde los cimientos. Pero antes que nada debemos saber en qué consisten.

El libro funciona por sí solo y no presupone otros conocimientos en el lector. Pero se podría ampliar. Por ejemplo, se puede complementar con la lectura de algunos de los textos originales que aparecen citados con frecuencia. Se trata de clásicos fácilmente accesibles como las *Meditaciones* de Descartes, los *Tres diálogos* de Berkeley, la *Investigación sobre el conocimiento humano* de Hume o sus *Diálogos acerca de la religión natural*. También puede ser leído perfectamente sin tener estos textos a mano. Una vez terminado, el lector puede dirigirse a los clásicos y a otros materiales como manuales de lógica o escritos sobre ética.

¿Sobre qué debemos pensar?

Estas son algunas de las preguntas que cualquiera de nosotros puede formularse acerca de sí mismo: ¿Qué soy? ¿Qué es la conciencia? ¿Puedo sobrevivir a mi muerte corporal? ¿Tengo la certeza de que las experiencias y las sensaciones de los demás son como las mías? ¿Si no puedo compartir la experiencia de los demás, puedo comunicarme con ellos? ¿Actuamos siempre de forma egoísta? ¿Acaso soy una especie de títere que cree actuar libremente cuando en realidad sus actos están programados?

También hay preguntas acerca del mundo: ¿Por qué existe algo, y no la nada? ¿Cuál es la diferencia entre el pasado y el futuro? ¿Por qué la causalidad actúa siempre desde el pasado hacia el futuro, o tiene sentido pensar que el futuro puede influir sobre el pasado? ¿Por

qué hay regularidades en la naturaleza? ¿Presupone el mundo un Creador? Y si es así, ¿podemos comprender por qué lo creó (él, ella o ellos)?

Finalmente, también hay preguntas acerca de nosotros y el mundo: ¿Cómo podemos estar seguros de que el mundo es tal como creemos que es? ¿Qué es el conocimiento y cuánto poseemos? ¿Qué es lo que convierte un campo de investigación en una ciencia? (¿Es una ciencia el psicoanálisis? ¿Lo es la economía?) ¿De dónde procede nuestro conocimiento de los objetos abstractos, como los números? ¿De dónde procede nuestro conocimiento de los valores y los deberes? ¿Cómo podemos saber si nuestras opiniones son objetivas o meramente subjetivas?

Lo raro de estas preguntas no es sólo su aspecto desconcertante a primera vista, sino también que desafían cualquier intento de encontrar la solución por un procedimiento sencillo. Si alguien me pregunta cuándo hay marea alta, sé dónde tengo que buscar la respuesta. Puedo consultar tablas autorizadas sobre mareas; puede ser que tenga una idea aproximada de cómo se forman, y si falla todo lo demás puedo ir y medir yo mismo la subida y la bajada de la marea. Una pregunta como ésta se responde por medio de la experiencia: es una cuestión empírica. Se puede resolver mediante procedimientos convencionales, que consisten en realizar ciertas observaciones y mediciones, o bien en aplicar ciertas reglas que han sido contrastadas con la experiencia y se sabe que funcionan. Las preguntas del párrafo anterior no son de este tipo. Parecen requerir un mayor grado de reflexión. No se nos ocurre de forma inmediata dónde tenemos que buscar la respuesta. Tal vez no estemos del todo seguros de lo que queremos decir con ellas o de qué podría valer como respuesta. ¿Cómo podría saber, por ejemplo, si no soy después de todo un títere que cree actuar libremente cuando en realidad sus actos están programados? ¿Deberíamos preguntar a científicos especializados en el cerebro? Pero ¿cómo sabrían ellos lo que tenían

que buscar? ¿Cómo sabrían que lo han encontrado? Imaginen el titular de prensa: «Los neurofisiólogos descubren que los seres humanos no son títeres». ¿Cómo es posible?

Entonces, ¿de dónde surgen estas preguntas tan desconcertantes?

En una palabra, de la autorreflexión. La capacidad de reflexión sobre sí mismos acompaña siempre a los seres humanos. Incluso cuando actuamos guiados por el hábito, seguimos siendo capaces de reflexionar sobre dicho hábito. Es algo habitual para nosotros pensar sobre algo y luego reflexionar sobre lo que estamos pensando. Podemos preguntarnos a nosotros mismos (o bien puede ser otra persona quien nos haga la pregunta) si sabemos de lo que estamos hablando. Para responder a eso debemos reflexionar sobre la posición que hemos adoptado, sobre nuestra propia comprensión de lo que decimos, sobre nuestras fuentes de autoridad. Podemos comenzar a dudar si de verdad sabemos lo que queremos decir. Podemos preguntarnos si lo que decimos es «objetivamente» cierto, o sólo el resultado de la perspectiva que adoptamos, de nuestra forma de enfocar la situación. Al pensar sobre estas cuestiones, nos encontramos con categorías tales como el conocimiento, la objetividad, la verdad, y es posible que también queramos pensar acerca de ellas. En tal caso lo que hacemos es reflexionar sobre conceptos, procedimientos y creencias que habitualmente nos limitamos a usar. Observamos el andamiaje de nuestro pensamiento, hacemos ingeniería de conceptos.

Este momento de reflexión puede plantearse en el curso de un debate de lo más normal. Un historiador, por ejemplo, se verá obligado en algún momento a preguntar por el significado de «objetividad», «evidencia» o incluso «verdad» en el campo de la historia. Un cosmólogo deberá detenerse por un momento antes de resolver ecuaciones que incluyan la letra t , y preguntar qué significa, por ejemplo, el flujo del tiempo, la dirección del tiempo o bien el origen del tiempo. Pero en este momento, lo admitan o no, se

convierten en filósofos. Y se ponen a hacer algo que se puede hacer bien o mal. La cuestión es hacerlo bien.

¿Cómo se aprende filosofía? Una forma más adecuada de plantear la pregunta sería decir: ¿Cómo se pueden adquirir nuevas técnicas de pensamiento? En este caso se trata de pensar acerca de las estructuras básicas del pensamiento. Esto es algo que puede hacerse bien o mal, de modo inteligente o inepto, pero hacerlo bien no depende en primer lugar de la adquisición de un cuerpo de conocimientos. Es más parecido a tocar bien el piano. El «cómo» es tan importante como el «qué». La personalidad filosófica más famosa del mundo clásico, el Sócrates de los diálogos platónicos, no se vanagloriaba de lo mucho que sabía. Al contrario, se enorgullecía de ser el único que sabía cuán poco era lo que sabía (de nuevo, reflexión). Lo que de verdad hacía bien —se supone, pues existen divergencias sobre este punto— era mostrar la debilidad de las pretensiones de conocimiento de los demás. Pensar correctamente significa evitar las confusiones, detectar las ambigüedades, centrar la atención en cada cosa por separado, construir argumentos fiables, darse cuenta de las alternativas posibles, y esta clase de cosas.

En resumen: nuestras ideas y conceptos se pueden comparar con las lentes a través de las cuales vemos el mundo. En filosofía, el objeto de estudio es la lente en sí misma. El éxito no depende tanto de lo que podamos saber al término de la investigación como de lo que podemos hacer cuando la cuestión se pone difícil: cuando soplan los vientos de la argumentación y la confusión lo invade todo. Se tiene éxito cuando se toman en serio las implicaciones de las ideas.

¿De qué nos sirve?

Todo eso está muy bien, pero ¿por qué molestarnos? ¿De qué nos sirve? La reflexión no es lo que mueve el mundo. No nos da de comer ni hace que vuelen los aviones. ¿Por qué no dejamos a un lado los

problemas de la reflexión y seguimos adelante con lo demás? Me gustaría esbozar tres tipos de respuesta, de distinto nivel de abstracción: alto, medio y bajo.

La respuesta de alto nivel cuestiona la propia pregunta —una estrategia típicamente filosófica, porque implica subir un grado de reflexión—. ¿Qué queremos decir cuando preguntamos de qué nos sirve? La reflexión no da de comer, pero tampoco la arquitectura, la música, el arte, la historia o la literatura. Deseamos comprendernos a nosotros mismos, eso es todo. Es algo que deseamos por sí mismo, igual que un científico puro o un matemático puro desean comprender el origen del universo o la teoría de conjuntos, o un músico desea resolver un problema de armonía o de contrapunto. No tenemos la mirada puesta en ninguna aplicación práctica. Buena parte de la vida se nos va en proyectos como el de criar más cerdos para comprar más tierra, de modo que podamos criar más cerdos y podamos comprar más tierra... El tiempo que nos queda, sea para la música o las matemáticas, o para leer a Platón o a Jane Austen, es un tiempo que hay que cuidar. Es el tiempo que dedicamos a mimar nuestra salud mental. Y nuestra salud mental es algo bueno en sí mismo, como la salud física. Además, también existe una retribución en términos de placer. Cuando nuestra salud física es buena, disfrutamos haciendo ejercicio físico y cuando nuestra salud mental es buena, disfrutamos ejercitando la mente.

Ésta es una respuesta muy abstracta. El problema no es que sea errónea, sino que sólo resultará convincente para alguien que ya esté convencido a medias —alguien que no haya formulado la pregunta original en un tono de voz demasiado agresivo.

Vamos pues con una respuesta de nivel medio. La reflexión es importante porque forma un *continuo* con la práctica: lo que pensamos sobre las cosas que hacemos influye en nuestro modo de hacerlas, o incluso en si las hacemos o no. Puede influir en nuestras investigaciones, en nuestra actitud hacia la gente que hace las cosas

de un modo distinto a como las hacemos nosotros o, en fin, en el conjunto de nuestra vida. Por poner un ejemplo sencillo, si nuestras reflexiones nos llevaran a creer en una vida después de la muerte, quizás estuviéramos dispuestos a soportar ciertas persecuciones que preferiríamos evitar si nos convenciéramos —al igual que muchos filósofos— de que tal idea no tiene sentido. El fatalismo, es decir, la creencia en que el futuro está fijado de antemano, sean cuales sean nuestras acciones, es una creencia puramente filosófica, pero que es capaz de paralizar cualquier acción. Traducido a términos más políticos, también puede expresar una cierta aquiescencia con el bajo estatus social asignado a ciertos sectores de la sociedad, lo cual puede resultar ventajoso para la gente de estatus más elevado que propicia la situación.

Consideremos algunos ejemplos comunes en Occidente. La reflexión acerca de la naturaleza humana ha llevado a mucha gente a considerar que somos esencialmente egoístas. Sólo buscamos el beneficio propio, sin preocuparnos realmente por nadie más. Un acto en apariencia desinteresado oculta la expectativa de un beneficio futuro. El paradigma dominante en las ciencias sociales es el *homo economicus*, el hombre económico que mira por sus intereses, en abierta competencia con los demás. Ahora bien, si la gente se convence de que todo el mundo es así, las relaciones entre las personas cambian; se vuelven menos confiadas y cooperadoras, más suspicaces. Ello influye en su forma de actuar en relación con los demás, y se traduce en diversos costes. Les resultará más difícil, y en ocasiones imposible, poner en marcha empresas conjuntas: pueden quedar atrapadas en una «guerra de todos contra todos», de acuerdo con la memorable expresión del filósofo Thomas Hobbes (1588-1679). En el mercado, su miedo a ser estafadas les hará incurrir en elevados costes de transacción. Si estoy convencido de que «un contrato verbal no vale el papel sobre el que está escrito», me veré obligado a pagar abogados para que elaboren contratos que impongan

sanciones, y si no confío en que los abogados vayan a poner algo más de su parte que el mínimo indispensable para justificar el cobro de sus honorarios, tendré que hacer revisar los contratos por otros abogados, y así sucesivamente. Pero todo ello se puede basar en un error filosófico: interpretar las motivaciones humanas desde un conjunto equivocado de categorías y, por lo tanto, sin comprender su naturaleza. Puede que la gente sea capaz de preocuparse por los demás, o que al menos esté dispuesta a poner algo de su parte o a cumplir sus promesas. Tal vez si se propusiera una visión más optimista, la gente sería capaz de vivir de acuerdo con ella, en cuyo caso mejoraría su calidad de vida. Así pues, este momento de reflexión, destinado a poner en claro las categorías adecuadas para comprender las motivaciones humanas, es una tarea importante desde un punto de vista *práctico*. No se limita al campo de lo académico, sino que escapa de él.

Veamos ahora un ejemplo muy distinto. El astrónomo holandés Nicolás Copérnico (1473-1543) reflexionó sobre nuestro modo de captar el movimiento y se dio cuenta de que lo percibimos de forma perspectivista, es decir, que nuestra noción del movimiento de las cosas depende del lugar en que nos encontramos y, en particular, de si nosotros mismos estamos en movimiento. Cuando estamos en un tren o en un aeropuerto, a menudo caemos en la ilusión de que el tren o el avión de al lado se pone en marcha, para luego darnos cuenta con sorpresa de que somos nosotros quienes nos movemos. Sin embargo, en la época de Copérnico la vida cotidiana ofrecía menos ejemplos de ello. Del mismo modo, podría ser que los desplazamientos aparentes de las estrellas y de los planetas no se debieran a su propio movimiento, según parece a primera vista, sino a que nosotros, los observadores, nos movemos. Y finalmente resultó que era así. La reflexión sobre la naturaleza del conocimiento —lo que los filósofos llaman *epistemología*, del término griego *episteme*, que significa «conocimiento»— se encuentra, pues, en el origen del

primer avance espectacular de la ciencia moderna. Las reflexiones de Einstein sobre nuestro modo de concebir la simultaneidad entre dos acontecimientos siguieron el mismo camino. Einstein se dio cuenta de que el resultado de nuestras mediciones depende de nuestro propio desplazamiento en relación con los hechos que observamos, lo cual llevó finalmente a la teoría especial de la relatividad (y el propio Einstein reconoció la importancia que habían tenido algunos filósofos del pasado a la hora de hacerle tomar conciencia de las complejidades epistemológicas de esta clase de observaciones).

Como último ejemplo, me gustaría examinar un problema filosófico en el que se ven atrapadas muchas personas cuando piensan acerca de la relación entre la mente y el cuerpo. Es habitual partir de una separación estricta entre la mente y el cuerpo, como si fueran dos cosas enteramente distintas. Esto que puede parecer una cuestión de mero sentido común puede llegar a interferir en nuestra práctica cotidiana de forma bastante insidiosa. Por ejemplo, resulta difícil comprender cómo se relacionan entre sí estos dos objetos separados. Los médicos podrían ver casi como inevitable el fracaso de los tratamientos cuando se trata de estados físicos que responden a causas mentales o psicológicas. La posibilidad de que hurgar en la mente del paciente pudiera originar algún cambio en el complejo sistema físico en que consiste su organismo podría parecerles rayana en lo imposible. Después de todo, la ciencia de toda la vida nos enseña que se requieren causas físicas y químicas para obtener efectos físicos y químicos. De este modo llegamos a un *a priori*, la certeza incuestionable de que cierta clase de tratamientos (las drogas y el *electroshock*, por decir algo) son «correctos» por definición y que otros (tratar a los pacientes de una forma más humana, mediante el consejo y el análisis) son «incorrectos»: acientíficos, poco serios y destinados al fracaso. Sin embargo, esta certeza no se basa en la ciencia, sino en una falsa teoría filosófica, y cambia por completo si poseemos una mejor concepción filosófica de la relación entre la

mente y el cuerpo. Esta mejor concepción debería hacernos ver que no hay nada sorprendente en el hecho de que exista una relación entre la mente y el cuerpo. Es bien sabido, por ejemplo, que el hecho de pensar en ciertas cosas (algo mental) puede hacer que una persona se sonroje (algo físico). Pensar en un peligro futuro puede dar lugar a toda clase de cambios corporales: el corazón late con fuerza, se aprietan los puños, los músculos se contraen. Por extrapolación, nada nos impide pensar que un estado mental como el optimismo pueda tener efectos físicos, como la desaparición de ciertas marcas o incluso la remisión de un cáncer. Si estas cosas ocurren, o no, se convierte en una cuestión empírica. La certeza incuestionable de que no pueden suceder se revela como el resultado de una mala comprensión de la estructura del pensamiento o, en otras palabras, de una mala filosofía, y en este sentido se la puede considerar acientífica. Las actitudes y prácticas de los médicos se pueden beneficiar positivamente de este cambio de planteamiento.

Así pues, la respuesta de nivel medio nos recuerda que la reflexión forma un continuo con la práctica y que ésta puede cambiar para mejor o para peor en función de la validez de nuestras reflexiones. Vivimos en un determinado sistema de pensamiento, del mismo modo que vivimos en un determinado edificio, y si nuestro edificio intelectual resulta estrecho y opresivo, nos interesa saber qué otras estructuras son posibles.

La respuesta de nivel bajo se limita a sacar algo más de brillo a esta argumentación, ya no en los terrenos limpios y bien ordenados de la economía o de la física, sino abajo, en los sótanos, donde la vida humana es menos civilizada. Uno de los *Caprichos* de Goya lleva por título *El sueño de la razón produce monstruos*. Goya creía que buena parte de las locuras de la humanidad eran el resultado del «sueño de la razón». Siempre habrá gente dispuesta a decirnos qué es lo que queremos, cómo nos lo van a dar y en qué deberíamos creer. Las creencias son contagiosas, y se puede convencer a la gente de casi

cualquier cosa. Estamos típicamente convencidos de que nuestra forma de hacer las cosas, nuestras creencias, nuestra religión, nuestros políticos son mejores que los de los demás, o de que los derechos que nos ha otorgado nuestro Dios están por encima de los suyos, o de que la defensa de nuestros intereses exige maniobras defensivas o ataques preventivos contra ellos. En último término, son estas ideas las que hacen que las personas se maten unas a otras. Ideas sobre cómo son los demás, o sobre quiénes somos nosotros o sobre cómo defender nuestros intereses o nuestros derechos, son las que nos llevan a la guerra, nos convierten en opresores sin sentir apenas mala conciencia o incluso hacen que nos resignemos a ser nosotros mismos los oprimidos. Cuando estas creencias van acompañadas del sueño de la razón, el único antídoto es un despertar crítico. La reflexión nos permite dar un paso atrás y tal vez reconocer cuán ciega o desviada era nuestra anterior forma de ver las cosas, o descubrir por lo menos si existen argumentos en favor de ella o si es simplemente subjetiva. Hacer esto de forma adecuada significa, una vez más, hacer ingeniería de conceptos.

La reflexión puede ser vista como algo peligroso, ya que no hay forma de saber por adelantado a dónde nos puede llevar. Siempre habrá argumentos en contra de ella. Mucha gente se siente incómoda o incluso se indigna ante las preguntas filosóficas. Algunos porque temen que sus ideas pudieran salir peor paradas de lo que desearían si empezaran a pensar sobre ellas. Otros tal vez se sientan atraídos por las llamadas «políticas de la identidad», en otras palabras, la clase de identificación con una determinada tradición o identidad de grupo, nacional o étnica, que les induce a cerrar la puerta ante cualquier extraño que venga a cuestionarla. Cualquier crítica será eliminada de un plumazo: sus valores son «inconmensurables» con los valores de los demás. Tan sólo aquellos que pertenecen a su reducido círculo pueden comprenderlos. La gente se siente cómoda cuando se retira a los senderos trillados de la tradición y no se

preocupa demasiado por su estructura, sus orígenes o incluso por las críticas que puedan merecer. La reflexión abre una ancha avenida hacia la crítica, y los senderos de la tradición acostumbran a alejarse de ella. En este sentido, las ideologías se convierten en círculos cerrados, siempre a punto para responder con indignación ante la mente inquisidora.

La tradición filosófica de los últimos dos milenios ha sido enemiga de esta amable complacencia. Ha insistido en que no vale la pena vivir la vida si uno no la somete a examen. Ha insistido en el poder de la reflexión racional para eliminar los elementos negativos de nuestras prácticas, y reemplazarlos por otros más positivos. Ha identificado la autorreflexión crítica con la libertad, de acuerdo con la idea de que sólo desde una adecuada comprensión de nosotros mismos podemos controlar la dirección en la que queremos ir. Sólo cuando contemplamos con prudencia nuestra propia situación, y la contemplamos como un todo, podemos comenzar a pensar en cambiarla. Marx dijo que los filósofos anteriores se habían esforzado por comprender el mundo, cuando en realidad se trataba de transformarlo —uno de los comentarios más célebres y absurdos de todos los tiempos y del todo desmentido por su propia práctica intelectual—. Hubiera sido mejor añadir que si uno no comprende el mundo, poco sabrá cómo cambiarlo, por lo menos hacia mejor. Rosencrantz y Guildenstern admiten que no saben tocar la flauta y sin embargo pretenden manipular a Hamlet. Cuando actuamos sin comprender, es bien posible que el mundo se haga eco de la respuesta de Hamlet: «¿Acaso creéis que soy más fácil de tocar que una flauta?».

Hay tendencias académicas en nuestra época que reaccionan en contra de estas ideas. Hay gente que pone en duda la idea misma de verdad, o de razón, o la posibilidad de una reflexión desinteresada. En general, lo que hacen es mala filosofía, a menudo sin darse cuenta siquiera de ello: ingenieros de conceptos incapaces de esbozar un

plan, ya no digamos diseñar una estructura. Volveremos a tratar esta cuestión en varios puntos del libro, pero de momento puedo prometer que este volumen se alinea sin vergüenza al lado de la tradición y en contra de cualquier escepticismo moderno o posmoderno sobre el valor de la reflexión.

La leyenda completa de los grabados de Goya es «La imaginación, abandonada por la razón, produce monstruos imposibles: unida a ella, es la madre de las artes y la fuente de sus maravillas». Así es como deberíamos pensar.

CAPÍTULO

1

El conocimiento

Tal vez el pensamiento más inquietante que hemos tenido muchos de nosotros, a menudo en un momento temprano de la infancia, es que el mundo entero podría ser un sueño; que los escenarios habituales y los objetos de la vida cotidiana podrían ser fruto de nuestra fantasía. La realidad en la que vivimos podría ser una realidad virtual, surgida de nuestra mente o quizás introducida en ella por algún siniestro Otro. Por supuesto, tales pensamientos se van del mismo modo que vinieron. La mayoría de nosotros hacemos lo posible por quitárnoslos de encima. Pero ¿qué razones tenemos para hacerlo? ¿Cómo podemos saber si el mundo tal como lo vemos nosotros es el mundo tal como es? ¿De qué modo concebimos la relación entre apariencia y realidad: las cosas tal como creemos que son, en oposición a las cosas tal como son?

Perder el mundo

Podríamos decir: todo comenzó el 10 de noviembre de 1619.

En tal fecha, en la ciudad de Ulm, al sur de Alemania, el filósofo y matemático francés René Descartes (1596-1650) se encerró en una

habitación al calor de una estufa y tuvo una revelación seguida por varios sueños, que interpretó como el anuncio de la tarea de su vida: el despliegue de la única vía verdadera hacia el conocimiento. El primer paso en este camino consistía en deshacerse de todo lo que antes había dado por supuesto y empezar de nuevo desde los fundamentos.

Por supuesto, no es que todo comenzara en 1619, ya que Descartes no era el primero. Los problemas que se planteó Descartes son tan viejos como el pensamiento humano. Se trata de problemas relativos al yo y a su carácter mortal, a su capacidad para el conocimiento y a la naturaleza del mundo que le rodea; problemas relativos a la realidad y a la ilusión. Todos ellos aparecen ya planteados en los textos filosóficos más antiguos que poseemos, los Vedas hindúes, que se remontan al 1.500 a.C. aproximadamente. El brillante ensayista francés Montaigne pertenecía a la generación inmediatamente anterior a Descartes, y su lema era el título de uno de sus magníficos ensayos: *Que sais-je?* («¿qué es lo que sé?»).

Tampoco se enfrentó Descartes a su empresa con una mente del todo inocente: disponía de un profundo conocimiento de las filosofías predominantes en su época, gracias a las enseñanzas de sus maestros jesuitas. Pero en los tiempos de Descartes las cosas estaban cambiando. El astrónomo holandés Copérnico había descubierto el modelo heliocéntrico (el centro es el Sol) del sistema solar. Personajes como Galileo, entre otros, estaban sentando las bases para una ciencia «mecánica» de la naturaleza. Según esta concepción, las únicas sustancias presentes en el espacio serían materiales, compuestas por «átomos», y su movimiento se debería exclusivamente a un conjunto de fuerzas mecánicas que la ciencia llegaría algún día a descubrir. Tanto Copérnico como Galileo entraron en conflicto con los guardianes de la ortodoxia católica, la Inquisición, ya que su concepción científica ponía en peligro, a los ojos de mucha gente, el lugar que correspondía a los seres humanos en el cosmos. Si

no hay nada más allá de lo que nos dice la ciencia, ¿qué sucede con el alma y la libertad humanas y con la relación del hombre con Dios?

Descartes era inteligente. Inventó la notación algebraica estándar y dio su nombre a las coordenadas cartesianas, que nos permiten traducir las figuras geométricas en ecuaciones algebraicas. Fue uno de los líderes de la revolución científica y realizó avances fundamentales no sólo para las matemáticas, sino también para la física, en especial para la óptica. Pero Descartes era también un católico convencido, de modo que para él era una tarea de gran importancia mostrar que el mundo científico que se estaba desplegando —vasto, frío, inhumano y mecánico— dejaba espacio, a pesar de todo, para Dios y la libertad, y para el espíritu humano.

De ahí surgen las obras más importantes de su vida, que culminan con las *Meditaciones*, publicadas en 1641, «en las que se demuestra la existencia de Dios y la distinción entre el alma y el cuerpo», de acuerdo con el subtítulo. Sin embargo, el texto subsiguiente explica también que la intención de Descartes es rescatar la visión moderna del mundo de las acusaciones de ateísmo y materialismo. El mundo científico puede resultar menos amenazador de lo que se temía. Puede convertirse en un lugar seguro para los seres humanos. Y la forma de hacerlo seguro es reflexionar sobre los fundamentos del conocimiento. De modo que comenzamos por Descartes porque fue el primer gran filósofo que se enfrentó con las implicaciones de la visión moderna y científica del mundo. Comenzar por los filósofos medievales o por los griegos significa muchas veces comenzar tan lejos de donde nos hallamos ahora que ponernos en su piel requiere un esfuerzo de imaginación probablemente excesivo. Comparativamente, Descartes es uno de nosotros, o al menos eso esperamos.

Resulta peligroso parafrasear a un filósofo, en especial uno tan conciso como Descartes. Voy a presentar algunos de los temas centrales de las *Meditaciones*. Lo haré con el mismo espíritu que si se

tratara de las «jugadas más importantes» de un partido. Un estudio más detallado del texto revelaría otros puntos de interés; un estudio más cercano al contexto histórico revelaría todavía otros distintos. Pero los puntos que señalaré bastarán para iluminar la mayoría de las cuestiones centrales de la filosofía posterior.

El Genio Maligno

Hay seis *Meditaciones*. En la primera, Descartes introduce el «método de la duda». Parte del principio de que si quiere establecer algo «firme y permanente» en las ciencias, debe deshacerse primero de todas sus opiniones comunes y empezar otra vez desde los fundamentos.

El motivo es que Descartes ha descubierto que incluso sus sentidos le engañan, y «es prudente no confiar del todo en quienes nos han engañado, aunque sólo fuera una vez».¹ Sin embargo, se objeta a sí mismo, sólo los locos («que dicen que visten de púrpura, estando desnudos, o que tienen la cabeza de barro, o que son calabazas, o que están hechos de vidrio»: los locos eran indudablemente pintorescos en el siglo XVII) niegan la incuestionable evidencia de sus sentidos.

En respuesta a eso, nos recuerda el caso de los sueños, durante los cuales nos representamos objetos de forma tan convincente a como acostumbran a hacerlo nuestros sentidos, aunque no guardan relación alguna con la realidad.

Sin embargo, se objeta de nuevo a sí mismo, los sueños son como los cuadros. El pintor puede modificar las escenas, pero en último término representa objetos derivados de cosas «reales», aunque sólo sean los colores. Siguiendo un razonamiento parecido, dice Descartes, incluso si los objetos más familiares (nuestros ojos, cabeza, manos y demás) fueran imaginarios, dependerían necesariamente de objetos más simples y universales que tendrían

que ser reales.

Pero ¿qué objetos? Descartes piensa que «no hay nada de lo que antes juzgaba verdadero de lo que no pueda dudar ahora». Y llegados a este punto, supondré, pues, no que un Dios óptimo, fuente de la verdad, sino cierto Genio Maligno, tan sumamente astuto como poderoso, ha puesto toda su industria en engañarme: pensaré que el cielo, el aire, la tierra, los colores, las figuras, los sonidos y todas las cosas externas no son diferentes de los engaños de los sueños, y que por medio de ellas ha tendido trampas a mi credulidad.²

Aquí tenemos al Genio Maligno. Una vez que se ha planteado esta temible hipótesis, la única defensa posible es prevenirse atentamente contra cualquier creencia falsa. Descartes reconoce que es algo difícil de llevar a la práctica, y que «cierta desidia» le hace volver a la vida normal, aunque intelectualmente no le queda otra opción que avanzar por las «inextricables tinieblas» de los problemas recién planteados. Así termina la primera *Meditación*.

Cogito, ergo sum

Al comienzo de la segunda *Meditación* encontramos a un Descartes abrumado por estas dudas. En beneficio de la investigación parte del supuesto de que «no tengo sentidos ni cuerpo». Pero:

¿Me he convencido también de que yo no soy? Ahora bien, si de algo me he convencido, ciertamente yo era. Pero hay cierto engañador, sumamente poderoso y astuto, que, de industria, siempre me engaña. Ahora bien, si él me engaña, sin lugar a dudas yo también existo; y engañeme cuanto pueda, que nunca conseguirá que yo no sea nada mientras piense que soy algo. De manera que, habiéndolo sopesado todo exhaustivamente, hay que establecer finalmente que esta proposición, *Yo soy, yo existo*, es necesariamente verdadera cada vez que la profiero o que la concibo.³

Este es el famoso *Cogito, ergo sum*: «Pienso, luego existo».

Habiendo salvado su «yo» de la marea general del escepticismo, Descartes pregunta ahora qué es este yo. Así como antes pensaba que conocía su cuerpo, y pensaba sobre sí mismo por analogía con su cuerpo, ahora se ve forzado a reconocer que el conocimiento de su yo no se basa en el conocimiento de su existencia corporal. En particular, va a tener problemas a la hora de imaginarlo. La imaginación tiene que ver con la contemplación de la forma o la imagen de un objeto corpóreo (un cuerpo humano o algo que ocupe un lugar en el espacio). Pero en este punto no sabemos nada acerca de objetos corpóreos. De modo que resulta inadecuado «imaginar» el yo a base de imaginar un sólido ente corporal que sea gordo o delgado, alto o bajo, tal como pueda verlo en un espejo.

Entonces, ¿en qué se basa este conocimiento del yo?

¿Pensar? Eso es: el pensamiento; esto es lo único que no puede separarse de mí. Yo soy, yo existo; es cierto. Pero ¿durante cuánto tiempo? Ciertamente, mientras pienso; pues tal vez podría suceder que si dejara de pensar completamente, al punto dejaría de ser. [...] Así pues, hablando con precisión, soy sólo una cosa pensante.⁴

La investigación toma ahora un camino ligeramente distinto. Descartes reconoce que una concepción del propio yo como algo corporal, que habita en un mundo extenso y espacial de objetos físicos, se impondrá otra vez de modo casi irresistible. Y se da cuenta de que el «yo» con el que se ha quedado es casi inaparente: «Ese no sé qué mío, que no es comprendido por la imaginación». Por lo tanto, «consideremos las cosas que vulgarmente se cree que se comprenden más distintamente que todas las demás, a saber, los cuerpos que tocamos y que vemos». Descartes considera en primer lugar el caso de una bola de cera. Tiene sabor y olor, color, figura y tamaño: todo ello es «manifiesto». Si la golpeas, emite un sonido. Pero luego acerca la cera al fuego y fíjense:

[L]os restos del sabor desaparecen, se disipa el olor, cambia el color, pierde la figura, crece la magnitud, se hace líquida, caliente, apenas se puede tocar, y si la golpeas ya no emitirá sonido. ¿Permanece aún la misma cera? Hay que reconocer que sí; nadie lo niega, nadie piensa otra cosa. ¿Qué habría entonces en ella que se comprendiera tan distintamente? Ciertamente, ninguna de las cosas que yo alcanzaba con los sentidos, pues todo lo que se refería al gusto, al olor, a la vista, al tacto, o al oído, ha cambiado ahora: la cera permanece.⁵

Descartes interpreta el resultado de este ejemplo como la demostración de que existe una percepción de la cera que es «sólo inspección de la mente», la cual puede llegar a ser «clara y distinta» si se es cuidadoso a la hora de concentrar la atención en aquello en que consiste la cera. De modo que, hacia el final de la segunda *Meditación*, concluye:

[P]uesto que ya sé que los cuerpos no son percibidos propiamente por los sentidos o por la facultad de imaginar, sino sólo por el entendimiento, y que no se perciben al tocarlos o al verlos, sino sólo porque se entienden, conozco claramente que no puedo percibir nada más fácil y evidentemente que mi propia mente.⁶

Fundamentos, interrogantes

¿Cómo deberíamos interpretar una propuesta filosófica como ésta? Al comienzo vemos cómo Descartes se esfuerza por fundamentar su método de duda extrema (también conocido como duda cartesiana, o tal como él mismo la llama, «hiperbólica», es decir, duda excesiva o exagerada). Pero ¿son satisfactorios sus argumentos? ¿En qué está pensando exactamente? Puede que sea esto:

Algunas veces los sentidos nos engañan. *Luego* por lo que sabemos, nos engañan siempre.

Pero se trata de un mal argumento: una falacia. Comparemos:

Algunas veces los periódicos cometen errores. *Luego* por lo que sabemos, cometen errores siempre.

El punto de partida o premisa es verdadero, pero la conclusión resulta harto improbable. E incluso se pueden encontrar ejemplos del argumento donde la premisa es verdadera, pero es imposible que la conclusión sea verdadera.

Algunos billetes son falsos. *Luego* por lo que sabemos, todos son falsos.

En este caso la conclusión es imposible, puesto que la idea misma de falsificación presupone la existencia de billetes o monedas válidos. Las falsificaciones son parasitarias de lo real. Los falsificadores necesitan billetes y monedas originales para hacer sus copias.

Un argumento es válido cuando no hay forma —en el sentido de que no hay forma posible— de que las premisas, o puntos de partida, sean verdaderas sin que la conclusión también lo sea (nos ocuparemos de esto con más detalle en el capítulo 6). Es verdadero si es válido y sus premisas son verdaderas, en cuyo caso la conclusión también es verdadera. El argumento que acabo de presentar es claramente inválido, puesto que no se distingue de otros ejemplos que nos llevan a conclusiones falsas desde premisas verdaderas. Pero ello mismo sugiere a su vez que resulta poco caritativo interpretar a Descartes en un sentido tan pobre. Deberíamos suponer que tiene en mente otra cosa, aunque lamentablemente no lo hace explícito. En este caso estaríamos buscando lo que se llama «una premisa oculta», es decir, un elemento necesario para que un argumento se sostenga, y que el autor quizás ha dado por supuesto, aunque no lo menciona de forma explícita. También podríamos reinterpretar a Descartes en el sentido de suponer una conclusión más débil. O quizás podemos hacer ambas cosas. El argumento podría ser:

Algunas veces los sentidos nos engañan. Somos incapaces de

distinguir cuándo lo hacen y cuándo no. *Luego* por lo que sabemos, cualquier experiencia sensorial en particular puede ser engañosa.

Este argumento parece tener más visos de validez. Si lo aplicamos al caso de los billetes y las falsificaciones, veremos que la conclusión parece seguirse. Pero la conclusión se refiere a cualquier experiencia en particular. Ya no se trata de que toda nuestra experiencia (en bloque, por decirlo así) pueda ser engañosa. Es la misma diferencia que hay entre «por lo que sabemos, cualquier billete en particular puede ser una falsificación» y «por lo que sabemos, todos los billetes son falsificaciones». La primera puede ser verdadera, mientras que la segunda no.

Sin embargo, quizás a esta altura de las *Meditaciones* Descartes no necesita más que la conclusión débil. Pero deberíamos fijarnos también en la segunda premisa de la versión refinada del argumento. ¿Es verdadera? ¿Es cierto que no podemos distinguir las veces que nos equivocamos —trátese de ilusiones, falsas apreciaciones o errores de interpretación sobre lo que vemos— de las veces que no? Para examinar esta cuestión me gustaría introducir una distinción. Tal vez sea cierto que no podemos detectar nuestros errores e ilusiones a primera vista. En eso consiste precisamente una ilusión. Pero ¿es cierto que tampoco podemos detectarlos con el paso del tiempo? Al contrario, más bien parece que sí somos capaces: podemos aprender, por ejemplo, a desconfiar de las imágenes de agua centelleante en el desierto, dado que son habitualmente ilusiones o espejismos, meros efectos producidos por la luz. Todavía más, el hecho de que en ocasiones seamos capaces de detectar los engaños de los sentidos viene seguramente presupuesto por el propio argumento de Descartes. ¿Por qué?

Porque Descartes presenta la primera premisa como un punto de partida, una verdad dada. Pero sólo podemos saber que los sentidos nos engañan en ocasiones si investigaciones ulteriores —a través de

los mismos sentidos— demuestran que lo han hecho. Descubrimos, por ejemplo, que un reflejo momentáneo de agua centelleante en el desierto nos ha llevado al error de pensar que allí había agua. Sin embargo, nos damos cuenta de nuestro error cuando nos acercamos y miramos con más atención, y si es preciso tocamos o escuchamos. De forma similar, el único motivo por el que sabemos, por ejemplo, que cualquier opinión inmediata e improvisada acerca del tamaño del Sol será errónea es que laboriosas observaciones ulteriores nos demuestran que el Sol es en realidad varias veces mayor que la Tierra. Por tanto, la segunda premisa sólo parece ser verdadera si tiene el sentido de que «no podemos distinguir a primera vista si nuestros sentidos nos están engañando». Ahora bien, para dar pie a las grandes dudas de Descartes parece que la premisa debería ser que «no podemos distinguir ni siquiera con tiempo y dedicación si nuestros sentidos nos están engañando o no». Y esta última no parece ser verdadera. En cierto sentido, podríamos decir que los sentidos son «autocorrectivos»: la propia experiencia sensible nos dice si en alguna ocasión la experiencia sensible nos ha inducido a error.

Acaso en previsión de esta clase de críticas, Descartes introduce en este punto la cuestión de los sueños. «Dentro» de un sueño vivimos experiencias que guardan cierto parecido con las experiencias de la vida ordinaria y, sin embargo, no hay nada en la realidad que corresponda al sueño. ¿Está diciendo Descartes que el conjunto de nuestra experiencia podría ser un sueño? Si fuera así, podríamos emplear de nuevo una distinción parecida a la que hemos planteado hace un momento: quizás no seamos capaces de distinguir inmediatamente o «a primera vista» si estamos soñando o no, pero en nuestra memoria no parece que tengamos ningún problema a la hora de distinguir en el pasado los sueños de los encuentros con la realidad.

Por otro lado, hay algo problemático en la idea de que toda

experiencia podría ser un sueño, pues ¿qué forma tenemos de saber si es verdad? A veces la gente se «da un pellizco» para asegurarse de que no está soñando. Pero ¿es realmente una buena prueba? ¿Acaso no podríamos soñar que el pellizco nos duele? Ciertamente, podemos esforzarnos por descubrir si nos hallamos en un sueño desde dentro del propio sueño. Sin embargo, por más que inventemos algún sutil experimento para saberlo, ¿acaso no podríamos soñar simplemente que realizamos el experimento o soñar que nos da como resultado que estamos despiertos?

Tal vez podríamos responder que los acontecimientos de la vida cotidiana poseen una amplitud y una coherencia que se halla ausente en los sueños. Los sueños son entrecortados y fragmentarios. No tienen pies ni cabeza. En cambio, la experiencia es vasta y majestuosa. Se despliega de un modo regular, o por lo menos pensamos que es así. A pesar de todo, en este caso Descartes tiene todo el derecho a preguntarse si no podría ser que la amplitud y la coherencia fueran ellas mismas engañosas. Eso le lleva al Genio Maligno, uno de los experimentos mentales más célebres de la historia de la filosofía. Su propósito es alertarnos sobre la posibilidad de que, al menos por lo que podemos saber, toda nuestra experiencia fuera como un sueño: del todo desconectada del mundo.

Es importante tener presentes dos cuestiones desde el principio. En primer lugar, Descartes es perfectamente consciente de que como seres humanos que vivimos y actuamos en el mundo no estamos en absoluto preocupados por tan peregrina posibilidad. De hecho, no seríamos capaces de hacerlo: tal como han señalado muchos filósofos, es psicológicamente imposible conservar alguna duda sobre el palpitante mundo exterior fuera del ámbito de la investigación. Pero eso no importa. Vale la pena hacer el esfuerzo de dudar en virtud de la tarea misma que ha emprendido Descartes. Esta tarea consiste en encontrar los fundamentos del conocimiento, en asegurar que sus creencias se han construido sobre un terreno sólido.

La investigación de Descartes se construye a partir de razonamientos puramente especulativos. En segundo lugar, Descartes no nos está pidiendo que creamos en la posible existencia del Genio Maligno. Tan sólo nos pide que la consideremos, como paso previo antes de encontrar el modo de descartarla. Es decir, Descartes piensa (no sin razón, ¿verdad?) que a menos que podamos descartar tal posibilidad, se mantiene en pie el reto del escepticismo: la posibilidad de que no poseamos ningún conocimiento y de que todas nuestras creencias sean completamente engañosas.

Podemos evaluar mejor el experimento mental si nos planteamos lo «real» que puede llegar a ser una realidad virtual. Veamos una variante actualizada del experimento. Imaginemos que un avance de la ciencia permitiera a un científico loco extraer nuestro cerebro y conservarlo luego en una cubeta llena de sustancias químicas que mantuvieran su funcionamiento normal.⁷ Imaginemos que el científico pudiera enviar señales a través de los canales normales de información (el nervio óptico y los nervios que transmiten las sensaciones acústicas, de tacto y de gusto). Si el científico tuviera buen corazón, nos daría la información como si el cerebro se hallara ubicado en un cuerpo normal y viviera una vida razonable: comer, jugar al golf o ver la televisión. Habría interacción, de modo que si, por ejemplo, enviáramos la señal de levantar la mano, recibiríamos una respuesta equivalente a levantar la mano. El científico nos habría introducido en una realidad virtual, de modo que nuestra mano virtual se levantaría. Y según parece, no tendríamos forma de saber lo que había sucedido, ya que para nosotros todo sería como si la vida normal continuara.

La versión cartesiana del experimento mental no menciona para nada los cerebros ni las cubetas. Pero si reflexionamos sobre ello nos daremos cuenta de que en realidad no los necesita para nada. Nuestras creencias sobre el cerebro y su papel en la producción de la experiencia consciente son creencias acerca del modo en que

funciona el mundo. ¡O sea que también podrían ser el resultado de las intromisiones del Genio Maligno! Quizás el Genio Maligno no tiene ninguna necesidad de ensuciarse las manos en cubetas. Simplemente nos envía las experiencias como sea que se envíen en la realidad real. Los cerebros y los nervios podrían no pertenecer más que a la realidad virtual.

Este experimento mental no se refiere a ilusiones o a sueños que tengan lugar realmente. Se limita a enfrentar el conjunto de nuestra experiencia con una realidad potencialmente extraña y perturbadora. Nótese también que no sirve de nada, en principio, señalar la amplitud y la coherencia de la experiencia cotidiana como argumento en contra de la hipótesis del Genio Maligno, pues no hay razón para que el Genio no haya de poder producir una experiencia tan coherente, amplia y extensa como desee.

Entonces, ¿cómo vamos a escapar a la hipótesis del Genio Maligno? Una vez que ha sido planteada, parece que somos impotentes frente a ella.

Y sin embargo, en medio de este mar de dudas, justo cuando todo parece sumido en la mayor de las oscuridades, Descartes descubre una roca sólida a la que agarrarse. *Cogito, ergo sum*: «Pienso, luego existo». (Sería mejor traducirlo por «Estoy pensando, luego existo». La premisa de Descartes no es «yo pienso» en el sentido de «yo esquío», lo cual podría ser cierto incluso si en aquel momento no estuviera esquiando. Corresponde más bien a «estoy esquiando».)

Incluso si lo que estoy experimentando es una realidad virtual, isigo siendo yo quien la experimenta! Y aparentemente no hay duda de que soy yo quien tiene estas experiencias o pensamientos (para Descartes «pensar» incluye «experimentar»).

¿Por qué se mantiene esta certeza? Veámoslo desde la perspectiva del Genio Maligno. Su intención era engañarme acerca de todo. Pero es lógicamente imposible que me haga creer que estoy pensando cuando en realidad no lo estoy haciendo. El Genio no puede convertir

en verdaderas al mismo tiempo estas dos afirmaciones:

Pienso que existo.

Me equivoco al pensar que existo.

La razón es que si la primera es verdadera, entonces tengo que existir para poder pensar. Por lo tanto, debo estar en lo cierto sobre la cuestión de si existo o no. En la medida en que pienso estas cosas (o incluso si pienso que las pienso), existo.

Puedo pensar que estoy esquiando cuando en realidad no lo estoy haciendo, sea porque estoy soñando o porque me engaña el Genio. Sin embargo, no puedo pensar que estoy pensando si no lo estoy haciendo. Esto es así porque en este caso (y sólo en este caso) el mero hecho de pensar que estoy pensando garantiza que estoy pensando. Eso mismo ya es pensar.

El escurridizo «yo»

Más allá del contexto de la duda, el «yo» que piensa es una persona que puede ser descrita de distintas formas. Por lo que respecta a mí, soy un profesor de filosofía de mediana edad, con una determinada personalidad, una historia, un entramado de relaciones sociales, una familia y demás. Pero en el contexto de la duda, todo esto desaparece: es parte de la realidad virtual. ¿Con qué «yo» nos hemos quedado entonces? Resulta bastante inaprensible: el sujeto puro del pensamiento. ¡Puede que ni siquiera tenga cuerpo! Todo esto nos lleva al siguiente paso.

Puedes escudriñar en tu propia mente, por así decirlo, en el intento de atrapar al «tú» esencial. Pero si tenemos en cuenta que este «tú» (o este «yo», desde tu punto de vista) se halla separado en estos momentos de toda marca normal de identidad (tu posición en el espacio, tu cuerpo, tus relaciones sociales, tu historia), parecería que no queda nada que atrapar. Eres consciente de tus propias

experiencias, pero según parece no del «yo» que es el sujeto de tales experiencias. También puedes tratar de imaginarte el «yo», hacerte una imagen de cómo es, por así decirlo. Pero tal como señala Descartes, la imaginación sirve ante todo para crear imágenes de objetos con una forma y un tamaño, y que ocupan un lugar en el espacio («objetos extensos»). El yo que resiste como la única roca sólida frente a la marea de la duda tal vez no sea un objeto extenso. La razón es que podemos estar seguros de su existencia cuando todavía mantenemos dudas acerca de los objetos extensos, ya que nos estamos tomando en serio la hipótesis del Genio Maligno.

Si tratamos de reconstruir el argumento en este punto, podemos suponer que Descartes está pensando lo siguiente:

No puedo dudar de mi existencia. Puedo dudar de la existencia de los objetos extensos (los «cuerpos»). Por lo tanto, no soy un cuerpo.

Dicho con pocas palabras: no dudamos del alma, dudamos del cuerpo, luego el alma es distinta del cuerpo. Si éste es el argumento de Descartes, entonces resulta plausible a primera vista, aunque también se puede considerar inválido. Examinemos un argumento parecido:

No puedo dudar de que estoy en esta habitación. Puedo dudar de si la persona que mañana recibirá malas noticias se halla en la habitación. Por lo tanto, mañana no voy a recibir malas noticias.

¡Bella prueba, y todavía mejor resultado! Esta clase de falacia recibe a menudo el nombre de «falacia del hombre enmascarado»: Sé quien es mi padre; no sé quien es el hombre enmascarado; *luego*, mi padre no es el hombre enmascarado.

Personalmente, dudo de que Descartes cayera realmente en esta falacia, al menos en esta *Meditación*. En este punto se interesa más bien por aquello que hace posible nuestro conocimiento del alma y

del cuerpo. No le interesa demostrar que son distintos, sino mostrar que el conocimiento del yo no depende del conocimiento del cuerpo, puesto que podemos tener la certeza de uno incluso cuando no la tenemos acerca del otro.

Un siglo más tarde, el filósofo alemán Georg Christoph Lichtenberg (1742-1799) señaló: «Deberíamos decir "piensa" igual que decimos "trueno". Incluso decir "cogito" es demasiado, si lo traducimos por "pienso"». ^{*18} (A Lichtenberg le gustaban los aforismos sentenciosos, y tuvo una importante influencia sobre otra figura posterior, Friedrich Nietzsche [1844-1900].)

La idea es que la referencia a un «yo», como si fuera una «cosa» o un sujeto del pensamiento, es ella misma una ilusión. No hay ningún «ello» que truene: lo único que podemos decir es que el tronar está sucediendo. Del mismo modo, lo que sugiere Lichtenberg es que Descartes no tiene derecho a un «yo» que piensa, al menos en el contexto de la duda. Propiamente no puede decir más que «un pensamiento está sucediendo».

La idea resulta bastante extraña. ¿Acaso es posible un pensamiento sin nadie que lo piense? Los pensamientos no van flotando por las habitaciones, a la espera, por decirlo así, de que alguien los atrape, como tampoco los dientes van flotando por ahí a la espera de encajar en una superficie adecuada. Volveremos sobre este punto en el capítulo 4. Pero entonces, ¿dónde está el error de Lichtenberg? Si Descartes no puede señalar ningún yo responsable del pensamiento, no tiene experiencia de él, no puede imaginarlo, ¿de dónde saca su certeza de que existe? Todavía más, ¿qué sentido tiene decir que existe?

Inteligentemente, Descartes deja a un lado este problema y plantea en su lugar una dificultad parecida en relación con «las cosas que vulgarmente se cree que se comprenden más distintamente que

^{1*} La idea queda más clara en inglés, puesto que el sujeto aparece en todas las oraciones. [N. del t.].

todas las demás», es decir, los cuerpos, los objetos que se encuentran en el espacio. Tal era la intención del ejemplo de la bola de cera. Veamos una posible reconstrucción del argumento:

En un momento determinado, mis sentidos me informan de la forma, el color, la dureza y el sabor que corresponden a la cera. Pero en otro momento mis sentidos me informan de que la forma, etc., que corresponden a la cera son distintos. Lo único que me muestran mis sentidos son estas distintas cualidades (a las que llamamos «cualidades sensibles», ya que las conocemos a través de los sentidos). A pesar de ello, emito un juicio de identidad: en ambos casos se trata de la misma cera. Luego, pertenece a la naturaleza de la cera el poseer diferentes cualidades sensibles en diferentes momentos. Luego, para saber en qué consiste la cera debo atender a mi intelecto, no a mis sentidos.

Cabe señalar que, si la reconstrucción del argumento es buena, Descartes no está negando que yo perciba la presencia de la cera en primer lugar gracias a los sentidos (asumiendo que hemos eliminado al Genio Maligno y que volvemos a confiar en nuestros sentidos). En realidad viene a decir algo muy parecido a esto. Lo que sugiere es que los sentidos son algo así como unos mensajeros que traen una información que precisa ser interpretada. Y esta interpretación, que consiste en identificar el objeto unitario a través de sus sucesivas apariencias, es tarea del intelecto, que lo consigue gracias al empleo de una serie de principios de clasificación o categorías, los cuales son a su vez susceptibles de ser investigados.

Por lo tanto, todo lo que sabemos acerca de la cera es que se trata de una «cosa» difícil de aprehender, capaz de adoptar distintas propiedades materiales (forma, tamaño, color o sabor). Y en el caso del «yo» nos encontramos con una «cosa» igualmente inaprensible que tiene pensamientos distintos en distintos momentos, de modo que quizá no deberíamos ver el yo como algo particularmente

misterioso, en comparación con objetos cotidianos como la cera. Quizá los yoes no sean más difíciles de comprender que los cuerpos, y sólo los vemos como algo distinto a causa de algún tipo de prejuicio. Volveremos a tratar la cuestión de la cera en el capítulo 7.

Ideas claras y distintas

Las dos primeras *Meditaciones* merecen un lugar entre los clásicos de la filosofía. Combinan profundidad, imaginación y rigor a un nivel que raras veces se ha igualado. Le dejan a uno sin aliento, pendiente de la continuación de la historia. Tenemos a Descartes agarrado a su minúscula y única roca, rodeado de un mar de dudas. Parece que se ha negado a sí mismo cualquier vía de escape. La vida todavía podría ser un sueño. Si volvemos a la metáfora de los fundamentos, diríamos que ha encontrado un lecho de roca sólida, pero que no dispone de materiales de construcción. La razón es que el propio listón que se ha fijado, el conocimiento «a prueba de Genios», parece negarle el empleo de los medios más naturales o «autoevidentes» de razonamiento si quiere defender que su conocimiento va más allá del Cogito. Nada impide suponer que el Genio nos podría estar engañando con falsos razonamientos. Nuestro intelecto podría ser todavía más falible que nuestros sentidos.

Curiosamente, no parece que Descartes lo vea del mismo modo. Lo que hace es reflexionar sobre el Cogito y preguntarse qué es lo que le confiere esta certeza tan especial. Se convence de que la razón es que percibe su verdad de un modo especialmente transparente, «claro y distinto». Se admite habitualmente que Descartes, que era matemático, tenía en mente un modelo matemático de claridad. Pensemos, por ejemplo, en una circunferencia. Imaginemos un diámetro y tracemos cuerdas desde sus dos extremos hacia un mismo punto de la circunferencia. Al encontrarse forman un ángulo recto. Si trazamos otras, veremos que al parecer siempre sucede lo

mismo. En este momento quizá no tenemos una idea demasiado clara de si existe algún motivo para ello. Pero ahora supongamos que ensayamos una demostración (trazar una línea desde el centro de la circunferencia hasta el vértice del ángulo recto del triángulo y resolver los dos triángulos resultantes). Después de hacerlo, nos damos cuenta de que el teorema debe verificarse. La comprensión llega como un «relámpago»: una intuición o una certeza cegadora de esta verdad geométrica en particular. Esto no es más que un ejemplo geométrico tomado al azar de un método que nos permite «ver» aquello que antes apenas alcanzábamos a comprender. ¡Ojalá fuéramos capaces de contemplar el resto de la realidad, la mente, el cuerpo, Dios, la libertad y la vida humana con la misma claridad y penetración! Pues bien, existe un ideal filosófico que afirma que sí somos capaces. Tal es el ideal del racionalismo: el poder de la razón pura, sin ayudas de ninguna clase. El racionalista sabe que las cosas son necesariamente de una manera y no de otra, sin necesidad de levantarse de su silla, igual que sucedía con el ángulo en el semicírculo. El conocimiento adquirido gracias a esta intuición racional recibe el nombre de *a priori*: nos damos cuenta de su verdad de forma inmediata, sin necesidad de acudir a la experiencia.

El argumento de la marca⁹

Con la confianza puesta en la claridad y la distinción, Descartes se permite un razonamiento. Al examinar su propio «yo», que es todo cuanto tiene en este momento, descubre que posee una idea de la perfección. Dicha idea, argumenta, supone una causa. Ahora bien, esta causa deberá poseer tanta «realidad» como la propia idea, y ello incluye la perfección. Ello implica que sólo puede tratarse de una causa perfecta, es decir, Dios. Luego, Dios existe, y ha dejado la idea de perfección en nuestras mentes como marca de su obra, del mismo modo que el artesano estampa su marca en sus trabajos.

En cuanto Descartes descubre a Dios, el mar de las dudas se vacía de golpe. La razón es que si Dios es perfecto, no puede ser que nos engañe: engañar no es propio de un ser bondadoso, ya no digamos si es perfecto. Por tanto, si cumplimos adecuadamente con nuestra parte, podemos estar seguros de que no seremos víctimas de ninguna ilusión. El mundo será tal como nosotros lo pensemos. Cumplir adecuadamente con nuestra parte significa sobre todo confiar sólo en las ideas claras y distintas.

¿Cómo debemos tomarnos el argumento de la «marca»? Veamos esta reconstrucción:

Tengo la idea de un ser perfecto. Tal idea debe tener una causa. La causa debe ser al menos tan perfecta como su efecto. Luego algo al menos tan perfecto como mi idea ha sido su causa. Por lo tanto tal objeto existe. Pero tal objeto debe ser perfecto, es decir, Dios.

Supongamos que le concedemos a Descartes la idea que menciona en la primera premisa. (Hay tradiciones teológicas que ni siquiera concederían eso; dirían que la perfección divina desafía cualquier comprensión y que, por lo tanto, no tenemos ninguna idea de él o de ello.) Ahora bien, ¿qué le da derecho a suponer que su idea debe tener una causa? ¿Acaso no puede haber hecho que simplemente no tengan ninguna causa? ¿Cosas que, por así decirlo, «pasan porque sí»? Después de todo, Descartes no puede apelar a ninguna experiencia normal ni científica desde su roca. Desde la desnuda soledad metafísica en la que se encuentra, ¿cómo puede negar que las cosas pasen porque sí? Y aunque pensara lo contrario, ¿acaso no debería preocuparle la posibilidad de que el Genio le estuviera manipulando y le hiciera pensar de este modo aunque fuera un error?

Sin embargo, la cosa empeora todavía más cuando llegamos al siguiente paso. Examinemos mi idea de una persona perfectamente puntual. ¿Acaso necesita una causa perfectamente puntual? Seguramente sería mejor plantearlo de otro modo. Puedo definir con

facilidad en qué consiste ser puntual en el caso de una persona: significa que nunca llega tarde (o quizá que nunca llega ni pronto ni tarde). Para comprender en qué consiste ser de esta manera no necesito conocer a nadie que lo sea. Puedo describir su comportamiento por adelantado. Sé cuál es la condición que deben cumplir sin necesidad de encontrarme con ninguno, o incluso de que alguna vez haya existido alguien así.

Es probable que Descartes rechazara la analogía. Su planteamiento sería más bien el siguiente: ¿puedo hacerme una idea de lo que sería un matemático perfecto? Para empezar, puedo pensar en un matemático que nunca cometa errores. Pero eso no acaba de ser lo que buscamos. Un matemático perfecto también debería tener imaginación e inventiva. Sin embargo, desde mis limitados conocimientos en matemáticas, tengo una idea bastante confusa de lo que eso significa. En general, me cuesta concebir o comprender los descubrimientos antes de que se produzcan: en caso contrario, los llevaría a cabo yo mismo. Así pues, tal vez sólo un matemático perfecto pueda darme una idea adecuada (una idea «clara y distinta») de cómo sería un matemático perfecto.

Es posible; pero en tal caso resulta hartamente dudoso que yo personalmente tenga una idea clara y distinta de un matemático perfecto y, por analogía, de un ser perfecto. Por lo general, lo que sucede es que cuando intento concebir esta idea pienso de una forma muy parecida a como lo hacía en el caso de una persona perfectamente puntual. Pienso en un agente que nunca comete errores, que no se comporta jamás de forma maliciosa, para quien nada es imposible, y así sucesivamente. Puede ser que añada en mi imaginación algo parecido a un aura, pero está claro que eso no nos aporta nada. Sin duda resulta presuntuoso, o incluso blasfemo, pretender que poseo una comprensión clara y completa de los atributos de Dios.

En realidad, en otro pasaje, el propio Descartes propone una

hermosa analogía que pone en peligro la base misma del argumento de la marca:

Podemos tocar una montaña con las manos, pero no podemos rodearla con los brazos tal como podríamos rodear un árbol o algún otro objeto que no fuera demasiado grande. Comprender algo significa abarcarlo con el propio pensamiento; para conocerlo basta con que nuestro pensamiento pueda tocarlo.¹⁰

Tal vez no seamos capaces de comprender las supuestas cualidades de Dios, por más que podamos acercarnos a ellas mediante definiciones. Pero en tal caso no podemos remitirnos a un ideal o arquetipo que nos permita comprenderlas.

La mayoría de nosotros coincidiremos, por lo tanto, en que el argumento de la marca está lejos de ser un argumento a prueba de Genios; tan lejos, en realidad, que resulta fácil resistirse a él incluso si no estamos bajo la presión de la duda extrema. En este punto, la historia de las ideas sugiere una serie de premisas ocultas que podrían servir para excusar a Descartes. Su optimismo acerca del argumento de la marca, sin duda superior al que podamos tener nosotros, se debía a ciertas ideas que había heredado de la tradición filosófica anterior. Una de las más importantes es que la causalidad propiamente dicha consiste en entregar algo al efecto, igual a como se entrega el testigo en una carrera de relevos. Así, por ejemplo, se requiere calor para calentar algo, o movimiento para poner algo en movimiento. Se trata de un principio que aparece una y otra vez en la historia de la filosofía y que volveremos a encontrar en más de una ocasión. En este caso, hizo pensar a Descartes que la «perfección» presente en su idea tenía que haber sido introducida en ella, por así decirlo, por una causa perfecta.

Pero es dudoso que tal principio esté a prueba de Genios. En realidad, ni siquiera es verdadero. Hoy nos resultan familiares algunas causas que no se parecen en nada a sus efectos. El movimiento de un

trozo de hierro en un campo magnético no se parece en nada a una comente eléctrica y, sin embargo, tal es su causa. De hecho, parece como si Descartes (una vez más influenciado por las ideas de la tradición filosófica anterior) hubiera ido a parar a la conclusión de que la idea de X comparte a su vez las propiedades de X. Luego, la idea de infinitud es una idea infinita. (¿La idea de algo sólido es una idea sólida?) De modo parecido, la idea de perfección sería una idea perfecta y requeriría una causa perfecta. Pero de nuevo podría ser el Genio quien hubiera introducido esta idea en nuestra mente y de nuevo no tenemos ningún motivo para caer en su trampa.

El círculo cartesiano

Descartes estaba convencido de que el argumento era correcto: todos sus pasos eran «claros y distintos». Así que ahora tiene a Dios, y no puede ser que Dios nos engañe. Pero recordemos que para llegar a este punto ha tenido que confiar en las ideas claras y distintas como fuentes de verdad. Y sin embargo, ¿no descubrimos un terrible fallo en este procedimiento? ¿Qué ha pasado con el Genio? ¿Acaso no podrían extraviarnos incluso nuestras ideas claras y distintas? Para salir al paso de esta posibilidad, según parece, Descartes se dirige a Dios —el mismo Dios cuya existencia acaba de demostrar— y lo convierte en garante de la verdad de aquello que percibimos como claro y distinto.

Fue uno de sus contemporáneos, Antoine Arnauld (1612-1694), quien dio la voz de alarma en este punto y acusó a Descartes de haber caído en un argumento circular, el infame «círculo cartesiano».¹¹ Parece que Descartes se ha comprometido con dos afirmaciones distintas. Consideremos el caso de que percibimos clara y distintamente una determinada *proposición p*, en cuyo caso *p* es verdadera. Podemos resumirlo con la expresión $(CDp \rightarrow Vp)$, que significa que si *p* es claro y distinto («CD»), entonces es verdadera

(«V»). Luego supongamos que «D» es el símbolo de «Dios existe y no nos engaña». El círculo consiste en que a veces Descartes parece defender la siguiente afirmación: puedo reconocer la verdad de $(CDp \rightarrow Vp)$ sólo si antes reconozco la verdad de D. Sin embargo, en otros puntos defiende esta otra afirmación: puedo reconocer la verdad de D sólo si antes reconozco la verdad de $(CDp \rightarrow Vp)$. Se parece al típico *impasse* de las mañanas, cuando nos haría falta un café para levantarnos de la cama y, sin embargo, debemos levantarnos de la cama para prepararnos el café.

Una de las dos tiene que ir primero. Se han dedicado muchas páginas a la cuestión de si Descartes cae o no cae realmente en esta trampa. Algunos comentaristas citan pasajes en los cuales parece que no sostiene realmente la primera. La idea es que D sólo sería necesaria para dar validez al recuerdo de las demostraciones. Así, en el momento en que percibimos algo de modo claro y distinto no nos hace falta confiar en ninguna otra cosa, ni siquiera en D, para afirmar su verdad. Sólo más tarde, cuando hemos olvidado ya la demostración, D se convierte en la única garantía de la validez de nuestro argumento anterior y de su verdad necesaria.

Otros comentaristas sugieren que Descartes no necesita la segunda. La existencia de Dios se le impone de forma clara y distinta, pero no necesita una regla general, del tipo $(CDp \rightarrow Vp)$, para dar validez a esta intuición. Puede estar seguro de la verdad de este resultado particular de la regla, sin necesidad de estar seguro de la regla en sí misma. Se trata de una sugerencia interesante, e introduce una importante verdad, a saber, que a menudo estamos más convencidos de los resultados particulares de nuestros juicios que de los principios generales a los que deberíamos apelar para defenderlos. Por ejemplo, puedo saber que una oración en particular es correcta desde el punto de vista gramatical, sin saber con seguridad cuál es la regla gramatical general que hace que sea correcta. Los filósofos han sido a menudo bastante reacios a esta

posibilidad. El admirado personaje de los *Diálogos* de Platón, Sócrates, exhibe con irritante orgullo su habilidad para provocar a uno de sus peles para que diga alguna cosa, para luego mostrar que no es capaz de defenderla por medio de principios generales, tras lo cual concluye que en realidad no tenía el menor derecho a realizar la primera afirmación. Sin embargo, el ejemplo de la regla gramatical sugiere que se trata de una mala inferencia. Consideremos también el caso de la percepción, que me permite reconocer una cosa determinada como un habitante de la Pomerania, un miembro de los Rolling Stones, o mi propia esposa, sin tener noción de ningún principio general que «justifique» mi veredicto. Mi sistema perceptivo debe actuar de acuerdo con una serie de principios generales o «algoritmos» para traducir los impulsos visuales en juicios, pero no tengo la menor idea de cuáles puedan ser, de modo que no podría responder a un Sócrates que me preguntara por los principios generales que justifican mi acto de reconocimiento. No haría más que vacilar y farfullar una u otra cosa. Y sin embargo ello no impide que reconozca al pomeranio, al Rolling Stone o a mi esposa. El método socrático sólo sirve para dar mala fama a los filósofos.

A pesar de todo, estamos obligados a preguntar por qué piensa Descartes que este resultado particular de la regla es indudable. ¿Por qué habríamos de considerar que su «intuición» clara y distinta de la existencia de Dios es un ejemplo también claro y distinto de intuición verdadera? Algunos de nosotros podemos abrigar la oscura sospecha de que el motivo es que la sola mención de Dios tiene el efecto de nublar la mente, en lugar de darle mayor claridad.

Para nuestros propósitos actuales, podemos abandonar la cuestión en este punto. La conclusión podría ser que existen indicios de una cierta dualidad de criterios. Descartes se ha apartado sin mucho ruido de la clase de escepticismo representado por el Genio Maligno en su esfuerzo por encontrar la manera de huir de la solitaria roca del Cogito. Ello podría sugerir que ha ido a parar a una isla desierta de la

que no hay escapatoria posible.

Fundamentos y redes

El gran filósofo escocés David Hume (1711-1776) criticó a Descartes del siguiente modo:

Hay una clase de escepticismo previo a todo estudio y filosofía, muy recomendado por Descartes y otros como una excelente salvaguardia contra el error y el juicio precipitado. Aconseja una duda universal, no sólo de nuestras opiniones y principios anteriores, sino también de nuestras mismas facultades de cuya veracidad, dicen ellos, nos hemos de asegurar por una cadena de razonamientos deducida a partir de algún principio original, que no puede ser falaz o engañoso. Pero ni hay tal principio original que tiene prerrogativa sobre todos los demás, que son evidentes por sí mismos y convincentes, o si lo hubiera, ¿podríamos dar un paso más allá de él si no fuese por el uso de esas mismas facultades en las que se supone que ya no tenemos confianza? La duda cartesiana, por tanto, si pudiera ser observada por criatura alguna (como claramente no lo es), sería absolutamente incurable y ningún razonamiento nos podría llevar jamás a un estado de seguridad y convicción sobre tema ninguno.¹²

Si el proyecto de Descartes consiste en proteger la veracidad de la razón frente a la duda universal a través de la propia razón, entonces está abocado al fracaso.

El desafío de Hume resulta convincente. Parece como si Descartes estuviera condenado al fracaso. ¿Y cuál sería el resultado si eso fuera cierto? ¿Un pesimismo general, en el sentido de un pesimismo acerca de la posibilidad de que haya alguna armonía entre las cosas tal como creemos que son y tal como son realmente? ¿Quizás algo distinto? Las otras posibilidades reclaman algunas explicaciones previas.

Una forma de pensar —la de Hume— asume el principio de que nuestro sistema de creencias requiere algún tipo de fundamento. Sin

embargo, niega que tal fundamento pueda tener la clase de estatus racional que exigía Descartes. La garantía de la veracidad de nuestros sentidos y de nuestros razonamientos es una parte inseparable de este fundamento. Dicha veracidad no puede ser demostrada a su vez a partir de ningún otro «principio original». Más allá de la especulación filosófica, todos confiamos de forma natural en nuestra experiencia común. Hemos sido educados en ella y a lo largo de este proceso hemos ido aprendiendo a reconocer las zonas de peligro (ilusiones, espejismos) sobre el fondo de creencias naturales que todos nos formamos. Para ello no necesitamos más que aquella naturaleza autocorrectora de nuestros sistemas de creencias que antes mencionábamos. Podríamos considerar este planteamiento como un fundacionalismo natural, no racional. (Lo cual no significa, por supuesto, que tenga nada de irracional. Se trata simplemente de que tal fundamento no cumple con los «criterios de racionalidad a prueba de genios» a los que aspiraba Descartes.) Este abandono de cualquier referencia a la razón fue defendido por el propio Hume a través de una serie de argumentos, algunos de los cuales serán tratados en su debido momento.

Este énfasis en las vías naturales de formación de la creencia se corresponde con otra actitud que comparten Hume y otros filósofos británicos de los siglos XVII y XVIII, a saber, su desconfianza ante la razón abandonada a sus propios medios. Para este grupo de filósofos, el mejor contacto entre la mente y el mundo no se produce en el momento en que cristaliza una prueba matemática, sino más bien cuando vemos y tocamos un objeto familiar. Su paradigma era el conocimiento derivado de la experiencia sensible, antes que de la razón. Por este motivo se les conoce como *empiristas*, mientras que Descartes sería un *racionalista* emblemático. Con todo, estas etiquetas ignoran un buen número de detalles relevantes. Por ejemplo, hay veces en las que el propio Descartes, cuando se ve presionado, parece decir que lo realmente bueno de las ideas claras y

distintas es que no podemos dudar de ellas mientras las estamos pensando. No se puede decir que sea una justificación puramente racional, ya que apela a la misma clase de impulso natural que Hume atribuye a las creencias empíricas más básicas. Por otro lado, pronto veremos cómo el primero de los empiristas británicos, John Locke (1632-1704), se comporta en cierto momento como el más racionalista de los racionalistas. Los grandes filósofos tienen la incómoda costumbre de resistirse a las etiquetas.

Visto desde esta perspectiva, el problema de Descartes es que confiaba demasiado en el poder de la razón. En lugar de ello, podemos apelar a la naturaleza, en este caso a nuestras tendencias naturales a la hora de formar nuestras creencias y corregirlas. ¿Y qué pasa con el Genio Maligno? Llegados a este punto, la auténtica moraleja que se puede extraer de los esfuerzos de Descartes es que cuando planteamos la pregunta de si nuestra experiencia y nuestra razón, en bloque, corresponden al mundo tal como es, también en bloque, será necesario un acto de fe para responder afirmativamente. «Dios» no es más que el nombre que le damos a lo que sea que garantiza esta armonía entre nuestras creencias y el mundo. Sin embargo, tal como señala Hume en el pasaje que acabamos de citar, no tenemos la menor necesidad de plantear esta pregunta en la vida normal. Tanto la duda hiperbólica como su respuesta serían, en este sentido, irreales.

Puede que esto suene razonable o, tal vez, meramente complaciente. Pero para levantar el cargo de complacencia podríamos señalar al menos lo siguiente: que consideremos irreal la duda no significa necesariamente que demos la espalda al problema de la armonía entre la apariencia y la realidad —cómo pensamos y cómo son las cosas—. Podemos acercarnos a él desde dentro del entramado normal de nuestras creencias. En realidad, cuando Hume ensayó esta aproximación quedó abrumado por las dificultades que se planteaban acerca de nuestra forma ordinaria de pensar sobre las

cosas: dificultades lo suficientemente graves como para introducir de nuevo el escepticismo sobre nuestra capacidad de conocer algo sobre el mundo. Este será el tema que trataremos en el capítulo 7.

Sin embargo, dos siglos más tarde aún nos queda un motivo para el optimismo. Podemos suponer que la evolución, presumiblemente responsable de los sentidos y las capacidades de razonamiento que poseemos, no los habría seleccionada (tal como los conocemos) si de hecho no funcionaran. Si la vista no nos informara, por ejemplo, de la presencia de depredadores, comida o posibles compañeros justo cuando estamos cerca de los depredadores, la comida o los posibles compañeros, no tendría la menor utilidad. De modo que está diseñada para hacer bien esa clase de cosas. La armonía entre nuestra mente y el mundo se debe al hecho de que el mundo es precisamente el responsable de nuestra mente. La función de ésta es representar el mundo de modo que podamos satisfacer nuestras necesidades; si estuviera diseñada de modo que no lo representara verazmente, no habríamos podido sobrevivir. Este argumento no está pensado para hacer frente a la hipótesis del Genio Maligno, ya que apela a las cosas que creemos saber sobre el mundo. Por desgracia, en su momento nos veremos obligados a pasar revista a las dudas de Hume, donde todo aquello que creemos saber acerca del mundo no sirve más que para lanzar nuevas dudas sobre este mismo conocimiento. Una respuesta relativamente distinta ignora la necesidad de buscar alguna clase de «fundamento», sea procedente de la razón, como pretendía Descartes, o meramente natural, como quería Hume. Este planteamiento pone el énfasis en cambio en la *coherencia estructural* de nuestro sistema cotidiano de creencias: el sólido entramado que forman unas con otras, a diferencia de las experiencias puntuales o las creencias procedentes de los sueños, que tienen un carácter fragmentario e incoherente. También señala un rasgo interesante de las estructuras coherentes, a saber, que no necesitan fundamentos. Una embarcación o una tela de araña están

hechas a partir de un entramado de partes que se relacionan de una cierta forma, y su resistencia se deriva precisamente de esta interrelación. No precisa una «base» o un «punto de partida» o un «fundamento». Cada una de las partes de esta clase de estructuras es sostenida por otras partes de la misma, sin que quede ninguna parte que sostenga a todas las demás y no esté ella misma sostenida por nada. De modo parecido, si una creencia cualquiera fuera puesta en duda, sería apoyada por otras creencias, a menos, evidentemente, que se descubriese que no había nada que la sostuviera, en cuyo caso debería ser abandonada. El filósofo austríaco Otto Neurath (1882-1945) propuso esta magnífica metáfora en referencia al conjunto de nuestro saber:

Somos como unos navegantes que se vieran obligados a reconstruir su barco en plena travesía, y que nunca tuvieran oportunidad de empezar de nuevo desde el bastimento.¹³

Cualquier parte puede ser reemplazada, a condición de que haya un número suficiente de las demás como para que el conjunto se sostenga. Pero no podemos cuestionar la estructura como tal, en bloque, y si nos atreviéramos a hacerlo iríamos a parar a la solitaria roca de Descartes.

Este planteamiento recibe habitualmente el nombre de «coherentismo». Su lema es que si bien todo argumento necesita alguna premisa, no hay nada que sea la premisa de todos los argumentos. No hay ningún fundamento que sostenga todo lo demás. En ciertos aspectos, el coherentismo resulta atractivo, pero en otros se demuestra insatisfactorio. Resulta atractivo en la medida en que nos permite desembarazarnos del inalcanzable fundamento. Ahora bien, no está claro que nos ofrezca a cambio lo suficiente como para reemplazarlo. El motivo es que todavía sigue en pie la posibilidad del Genio Maligno, es decir, la posibilidad de que nuestro sistema de creencias sea exhaustivo, coherente y bien trabado, pero

completamente erróneo. Tal como señalé en la introducción de este capítulo, incluso los niños se preguntan de forma espontánea si es posible que toda nuestra experiencia sea un sueño. Quizás estaríamos de acuerdo con Descartes en que si las opciones son el coherentismo o bien el escepticismo, la más honesta sería el escepticismo.

Valdrá la pena recordar las cuatro opciones existentes en epistemología (o teoría del conocimiento). Hay un fundacionalismo racional, en la línea de Descartes. Hay un fundacionalismo natural, en la línea de Hume. Por otro lado está el coherentismo. Y sobre todos ellos se cierne el escepticismo, es decir, la idea de que no es posible el conocimiento. Cada una de ellas tiene destacados defensores, de modo que sea cual sea la que escoja el lector, contará con una buena compañía filosófica. Puede pensarse que Descartes acertó en casi todo, o bien que se equivocó en casi todo. Lo extraño del caso es que cualquiera de las respuestas resulta igualmente válida.

Escepticismos locales

El escepticismo no toma tan sólo el aspecto global que le dio Descartes, sino que puede quedar restringido a un campo en particular. Alguien puede estar convencido, por ejemplo, de la validez del conocimiento científico y mantener, sin embargo, serias dudas acerca de la validez del conocimiento en el campo de la ética, la política o la crítica literaria. En estos casos el motivo de nuestra inseguridad no es una duda hiperbólica, sino tan sólo un poco de cautela. Sin embargo, cuando el campo tiene cierta generalidad, el escepticismo resulta a veces de lo más desconcertante. El filósofo Bertrand Russell (1872-1970) dedicó algunas consideraciones al problema del tiempo.¹⁴ ¿Cómo puedo saber que el mundo no comenzó a existir hace tan sólo escasos momentos, aunque plagado de falsos vestigios de una era mucho más antigua? Estos vestigios incluirían,

por supuesto, las alteraciones en el cerebro que percibimos como recuerdos y, en general, todo aquello que interpretamos como señales de épocas pretéritas. En realidad, los pensadores Victorianos sugirieron algo muy parecido en el campo de la geología cuando trataban de reconciliar la explicación bíblica de la historia del mundo con la evidencia fósil. De acuerdo con su hipótesis, hace unos 4.000 años Dios habría creado todos los falsos indicios de que la Tierra tenía aproximadamente 4.000 millones de años de antigüedad (hoy podríamos añadir los falsos indicios de que el universo tiene 13.000 millones de años de antigüedad). El argumento nunca fue demasiado popular, probablemente porque si uno es escéptico en relación con el tiempo, pronto lo será en relación con todo lo demás, o quizás porque presenta a Dios como algo muy parecido a un gracioso a gran escala. La hipótesis de Russell resulta casi tan rebuscada como el Genio Maligno de Descartes.

Sin embargo, hay algo altamente intrigante en el escenario imaginado por Russell. Y es que, desde un punto de vista científico, iresulta mucho más *probable* que la alternativa en la que todos creemos! El motivo es que la ciencia afirma que la baja entropía, en otras palabras, los sistemas altamente ordenados son más improbables.¹⁵ Por si fuera poco, a medida que el cosmos evoluciona aumenta también la entropía, o el desorden. El humo no regresa jamás al cigarrillo; la pasta de dientes nunca se vuelve a introducir en el tubo. Lo extraordinario es que hubiera alguna vez suficiente orden en el mundo como para que el humo se hallara aún en el cigarrillo o la pasta de dientes en el tubo. Se podría defender, pues, que es más fácil que exista un mundo moderadamente desordenado como el que conocemos hoy, que un predecesor suyo más ordenado y con menor entropía. Intuitivamente parece que hay más posibilidades de que sea así, del mismo modo que hay más posibilidades de construir palabras de cuatro o cinco letras al comienzo de una partida de Scrabble, que una de siete letras. Lo más probable es que primero salga una

palabra de cuatro letras y luego otra de siete. De modo parecido, prosigue el argumento, parece como si a Dios o a la naturaleza le fuera a costar menos trabajo crear de la nada el mundo tal como es hoy, que crear de la nada el mundo de baja entropía que se supone existió hace algo así como 13.000 millones de años. Por lo tanto, es más probable que sucediera de este modo. En una competición abierta de probabilidades entre la rebuscada hipótesis de Russell y el sentido común, gana Russell. Dejo la cuestión a la consideración del lector.

La moral

¿Cuál habría de ser, pues, nuestra opinión acerca del conocimiento? El conocimiento supone autoridad: es mejor escuchar a aquellos que saben más. También supone fiabilidad: aquellos que más saben son también más fiables a la hora de captar la verdad, como los buenos instrumentos. Afirmar que uno sabe algo significa afirmar que uno confía en la propia fiabilidad. Y reconocerle autoridad a alguien o a algún método supone considerarlo fiable. Los inquietantes escenarios planteados por Descartes o Russell ponen en cuestión esta confianza en la propia fiabilidad. Una vez que se han planteado sus rebuscadas hipótesis, se hace más difícil pensar en una conexión fiable entre las cosas tal como creemos que son y tal como son realmente. Acaso podríamos recuperar esta confianza si fuéramos capaces de demostrar que tales hipótesis son imposibles, o que por lo menos no tienen ninguna posibilidad real de darse en un mundo como el nuestro. El problema es que resulta difícil decir por qué son imposibles y, por otro lado, nuestro sentido de las probabilidades pierde pie cuando llegamos al reino de la abstracción pura. De modo que resulta difícil argumentar que no tienen ninguna opción de ser verdaderas, a menos que confiemos en las mismas opiniones que ellas ponen en cuestión. Por ello, el escepticismo nunca

deja de ejercer su atracción, o bien de amenazarnos. Quizás estemos explorando el mundo de forma fiable, pero quizás no. Para volver a la analogía de la ingeniería que utilicé en la «Introducción», parece que la estructura de nuestro pensamiento tiene brechas importantes: en este caso, la brecha que separa la apariencia de las cosas de lo que son realmente. Nos concedemos a nosotros mismos el derecho a saltarnos estas brechas. Pero si lo hacemos sin antes encontrar un buen indicio de nuestra propia fiabilidad, es decir, de nuestra sintonía con la verdad, tal derecho parece que no tiene ningún fundamento. Y éste es el punto sobre el que insiste el escéptico. Cualquier confianza que tengamos en una posible armonía entre las cosas tal como creemos que son y tal como son realmente le parecerá un puro acto de fe.

Descartes nos legó un problema para el conocimiento. También nos dejó con serias dificultades a la hora de comprender el lugar que ocupan nuestras mentes en la naturaleza.

Y finalmente la revolución científica en la que ocupaba un lugar tan distinguido trajo consigo graves problemas a la hora de comprender el mundo en el que nos encontramos. Hemos tenido ocasión de tratar el problema del conocimiento. En el próximo capítulo nos ocuparemos del problema de la mente.

CAPÍTULO

2

La mente

Supongamos que dejamos a un lado el problema general de la armonía entre nuestra forma de ver el mundo y la forma en que el mundo es realmente. Mantengamos los dedos cruzados y supongamos que de verdad sabemos lo que de modo natural creemos saber. Pero ¿hasta qué punto son consistentes nuestras creencias? Descartes convirtió el yo y la mente en objetos especiales, íntimos, susceptibles de un conocimiento inmediato. O tal vez cada uno de nosotros descubre en su propia mente este objeto especial e íntimo, susceptible de un conocimiento inmediato, pues aunque pueda encaramarme al Cogito en medio del mar de las dudas, me resulta imposible llegar hasta la naturaleza de «tu» mente. De modo que ¿cómo puedo llegar a saber algo de tu vida mental? ¿Cómo puedo saber, por ejemplo, que ves el color azul igual como lo veo yo? ¿No podría ser que algunos de nosotros sintiéramos más el dolor, pero nos quejáramos menos, o que otros lo sintieran menos, pero se quejaran más? ¿Qué debemos pensar acerca de la mente y el cuerpo, el cerebro y el comportamiento?

El fantasma en la máquina

Tal como hemos visto, Descartes empleó una estrategia que le llevó a la conclusión de que el conocimiento que tenemos de nuestra propia mente es más cierto y seguro que el del resto del mundo. No obstante, Descartes era también un científico. Sus descubrimientos en el campo de la óptica contribuyeron a sentar las bases de la disciplina. Practicó disecciones y alcanzó una razonable comprensión del proceso de transmisión de los impulsos a través de los nervios hasta el cerebro. Sabía que se trataba de una transmisión física, una «sacudida» o «movimiento violento» de los nervios, o tal como lo plantearíamos hoy, un impulso electroquímico transmitido a través del sistema nervioso. Los sentidos de la vista, el tacto, el gusto, el olfato y el oído activan el sistema nervioso, el cual transmite los mensajes al cerebro. Éste, por supuesto, no se reduce a una masa indiferenciada. Unas partes del cerebro transmiten señales a otras partes del cerebro y luego de vuelta al cuerpo: en el proceso intervienen complejos patrones de activación. Todo esto forma parte de la neurofisiología. Se trata de hechos que en principio pueden ser observados públicamente: con los instrumentos adecuados, se pueden mostrar los patrones de activación a toda una clase.¹

¿Y entonces qué?

Pues bien, entonces llega el momento mágico. La «mente» (la cosa pensante o *res cogitans*) se activa también, y se abre el mundo entero de la experiencia. El sujeto ve colores, oye ruidos, percibe texturas y temperaturas y tiene sensaciones de gusto y olfato. Este mundo de la experiencia se compone de estados mentales, es decir, estados que se dan en la conciencia subjetiva. Estos estados no se pueden observar públicamente. Son privados. Aunque toda la clase vea cómo se activan una serie de neuronas, sólo una persona sentirá el dolor. Descartes llegó incluso a identificar el lugar donde se produce el acontecimiento mágico. Basándose en una serie de motivos neurofisiológicos bastante razonables, pensó que el lugar donde se producía el paso de los mensajes desde el reino de la física

al reino de lo mental tenía que ser la glándula pineal, una estructura que se encuentra situada en el centro del cerebro.

En la concepción de Descartes, no se trata sólo de que los estados mentales sean distintos de los estados físicos. Pertenecen también a una clase distinta de sustancia, la sustancia inmaterial, una especie de elemento espiritual o ectoplasma. Hablando estrictamente, si digo «pensé en la Reina y saludé», estoy cayendo en una cierta ambigüedad: el «yo» que es el sujeto del pensar no es el mismo «yo» que saluda, que es el cuerpo. Los pensamientos y las experiencias son modificaciones de un tipo de sustancia; el movimiento y la posición pertenecen a la otra. Esta parte de la doctrina de Descartes lo convierte en un «dualista de las sustancias». No se trata sólo de que haya dos clases distintas de propiedades (las propiedades mentales y las físicas) y de que las personas puedan poseer ambas, se trata de que hay también dos clases de portadores de las propiedades. Sin duda, esto resulta muy oportuno desde el punto de vista teológico: deja la puerta abierta a la inmortalidad del alma, ya que no hay ningún motivo por el cual el elemento espiritual deba tener la misma esperanza de vida que algo del estilo de un cuerpo físico. Sin embargo, el dualismo de las sustancias no es inevitable. Se podría sostener que las propiedades mentales y las físicas son muy distintas y que sin embargo ambas corresponden a un mismo cuerpo organizado; después de todo, la masa y la velocidad son dos propiedades bien distintas y sin embargo ambas se aplican a los proyectiles. Aquellos que defienden que hay dos clases distintas de propiedades (las mentales y las físicas) pero que pueden corresponder a un elemento único (aquello de lo que están hechos los grandes animales, sea lo que sea) reciben el nombre de «dualistas de las propiedades».

Descartes nos acerca a una concepción que quedó perfectamente resumida por Gilbert Ryle (1900-1976) cuando dijo que convertía al ser humano en un «fantasma en una máquina».² Los hechos relativos

a la máquina, el cuerpo físico, son como los demás hechos que puedan producirse en el mundo físico. Consisten en una serie de interacciones entre elementos conocidos: moléculas y átomos, campos eléctricos y fuerzas. Los hechos que se producen en la parte espiritual, la mente, son del todo distintos. Quizá tienen lugar en alguna clase de elemento espiritual, sea ectoplasma o cualquier otro elemento inmaterial (no físico) del que estén hechos los espíritus y los ángeles. De acuerdo con la mentalidad popular, ambos funcionan sin necesidad de cuerpo físico. En cambio, en el ser humano normal hay una estrecha relación entre los hechos de una clase y los de otra: clavarle a alguien un alfiler produce ciertos cambios físicos, pero también causa el estado mental del dolor. Y viceversa: el acto mental de recordar una metedura de pata puede causar ciertos estados físicos como suspiros y rubores. De modo que los hechos que se producen en un reino influyen en los que se producen en el otro, aunque en principio se trata de dos reinos enteramente distintos.

Zombis y mutantes

Por supuesto, esta concepción no es un rasgo peculiar de Descartes, sino que se halla presupuesta en muchas de las religiones más importantes del mundo: forma parte de cualquier doctrina que defienda nuestra capacidad para sobrevivir a la muerte corporal o la posible separación del alma y el cuerpo. Sin embargo, se trata de una concepción que se enfrenta a problemas enormes, e incluso hay argumentos para pensar que insuperables.

La primera familia de problemas es epistemológica. Acabo de decir que en el ser humano normal existe una estrecha relación entre los estados mentales y los físicos. Pero ¿qué nos da derecho a creer eso? Una posibilidad es que las cosas fueran del siguiente modo:

La opción zombi. Los zombis se parecen a cualquiera de nosotros y se

comportan igual que cualquiera de nosotros. Su naturaleza física es indistinguible de la nuestra. Si abriéramos el cerebro de un zombi, veríamos que funciona exactamente igual que el de cualquiera de nosotros. Si pinchamos a un zombi, él o ella dirán «au» tal como haría cualquiera de nosotros. Sin embargo, los zombis no tienen conciencia. No hay ningún espíritu dentro de ellos.

Como los zombis tienen el mismo aspecto y se comportan igual que cualquiera de nosotros, no hay forma de distinguir quién de nosotros es un zombi y quién es consciente tal como podamos serlo tú o yo. En cualquier caso, tal como pueda serlo yo, pues ahora que he planteado la opción zombi, me doy cuenta de que no puedo estar realmente seguro de lo que seas tú o cualquier otro. Tal vez la conciencia sea un correlato extremadamente inusual del complejo sistema formado por el cerebro y el cuerpo. Tal vez yo sea el único caso: tal vez todos vosotros seáis zombis.

Otra posibilidad es que las cosas fueran de este otro modo:

La opción mutante. Los mutantes se parecen a cualquiera de nosotros y se comportan igual que cualquiera de nosotros. Su naturaleza física es indistinguible de la nuestra. Si abriéramos el cerebro de un mutante, veríamos que funciona exactamente igual que el de cualquiera de nosotros. Si pinchamos a un mutante, él o ella dirán «au» tal como haría cualquiera de nosotros.

A diferencia de los zombis, los mutantes tienen conciencia. Hay un espíritu dentro de ellos. Sin embargo, lo que sucede en el espíritu del mutante no es lo que nosotros esperaríamos. Cuando un mutante recibe un pinchazo, por ejemplo, quizá lo que experimenta es el estado mental correspondiente a oír un do medio tocado por un clarinete. Si dice «au» es porque su cerebro funciona igual que el nuestro y se comporta igual que nosotros, de modo que recibir el pinchazo de un alfiler pone en marcha una serie de procesos que comportan unos cambios que le llevan finalmente a decir «au», exactamente igual que todos nosotros. Quizá cuando oye realmente un do medio tocado por un clarinete siente un terrible dolor, aunque no haga más que sonreír beatíficamente. Un

mutante quizá vea los buzones de correos británicos de color amarillo; otro quizá vea los narcisos de color azul. Lo que sucede en la conciencia del mutante no guarda ninguna relación con lo que suceda en tu mente o en la mía. O en cualquier caso en la mía, pues ahora que he planteado la opción mutante, me doy cuenta de que no puedo estar realmente seguro de lo que seas tú o cualquier otro. Tal vez todos vosotros seáis mutantes, comparados conmigo.

La cuestión es que, desde la perspectiva del dualismo cartesiano, ambas opciones parecen estar completamente abiertas. Ambas resultan inquietantes, y no son algo que tengamos normalmente en cuenta (aunque sospecho que se nos pasan por la cabeza más a menudo que las peregrinas hipótesis del primer capítulo).

Una forma de enfrentarse a ellas es coger el toro por los cuernos. Podemos decir: de acuerdo, supongamos que ambas opciones están completamente abiertas. Tal vez nunca pueda saber cómo es en realidad la mente de otra persona, qué sucede en ella o incluso si posee algún tipo de vida mental. Pero ¿acaso no puedo suponer que la vida mental de las demás personas es muy parecida a la mía? ¿Acaso no es razonable tomarse a uno mismo como modelo para todos los demás? En este caso no se trataría tanto de un conocimiento como de una hipótesis o una conjetura, aunque podríamos decir que es una conjetura razonable. Esto se conoce como el argumento de la analogía para la existencia de las otras mentes.

Este argumento tiene el problema de que resulta increíblemente débil. Tal como preguntó despectivamente el gran filósofo austríaco Ludwig Wittgenstein (1889-1951): «¿Y cómo puedo generalizar ese único caso tan irresponsablemente?».³ El simple hecho de que, acaso por pura suerte, en un caso concreto —el mío propio— exista una vida mental de un determinado tipo, asociada a un cerebro y a un cuerpo, resulta una base muy débil para suponer que la misma asociación se da en todos los demás casos. Si tengo una caja y dentro

de ella hay un escarabajo, eso me da poca base para suponer que todo el mundo que tenga una caja tendrá también un escarabajo dentro de ella.

Peor todavía, me da poca base para negar que haya escarabajos en lugares distintos de las cajas. Acaso objetos físicamente muy diferentes de nosotros sean tan conscientes como pueda serlo yo: las rocas o las flores, por ejemplo.

Puede que nos sintamos inclinados a «quitarnos de encima» las opciones zombi y mutante. Podemos pensar que son ejemplos de fantasía filosófica, del todo irreales o en cualquier caso inverificables. Pero no es una reacción muy inteligente. A los neurofisiólogos, por ejemplo, les resulta imposible observar la experiencia consciente con las mismas técnicas que les permiten observar las neuronas y las sinapsis y los patrones de actividad cerebral; de acuerdo con el ejemplo anterior, no pueden mostrarla en pantalla para sus alumnos en la sala de conferencias. Pero entonces, desde la perspectiva del dualismo cartesiano, la opción en la que creemos todos de forma natural, a saber, que las demás personas no son zombis ni mutantes, ¡es ella misma inverificable! Se limita a ser un ciego artículo de fe. Alguien que defienda la opción zombi no se encuentra a estos efectos en peor posición que el resto de nosotros.

Nuestra concepción de la mente permite las opciones zombi y mutante, lo cual nos da derecho a suponer incluso que tienen bastantes probabilidades a su favor o, por lo menos, tantas como cualquier otra cosa. Puesto que si no es *a priori* falso que las demás personas sean zombis, ¿por qué habría de ser *a priori* más probable que fueran conscientes como yo?

¿Por qué hablan tanto los filósofos sobre extrañas posibilidades que las demás personas ignoran felizmente (algo que es en parte culpable del mal nombre y del aspecto esotérico de la disciplina)? El motivo es que tales posibilidades se usan para someter a examen una determinada concepción de las cosas. En este caso se usan para

poner a prueba la concepción de la mente y la materia que ha permitido plantearlas. El argumento es que si pensamos la mente y la materia a la manera cartesiana, entonces estarían perfectamente abiertas toda clase de extrañas posibilidades, sobre las que no podríamos decir nada. Luego, como esta situación es intolerable, deberíamos reconsiderar nuestra concepción de las cosas (esto se llama «metafísica»). Una mejor concepción de la mente y del lugar que ocupa en la naturaleza evitaría esta clase de posibilidades. La intención no es revolcarse en el escepticismo, sino retirarse de cualquier filosofía que deje alguna opción abierta al escepticismo. Sería como decir: de acuerdo con el dualismo cartesiano, tanto la opción zombi como la opción mutante están perfectamente abiertas. Pero ello no hace más que demostrar que algo falla en el dualismo cartesiano. Simplemente, lo mental y lo físico no son tan distintos como se pretende. Porque realmente no es posible que (por decir algo) alguien que haya recibido un golpe en un dedo del pie y esté gritando de dolor lo haga porque experimente un estado mental parecido al mío cuando oigo un do medio tocado por un clarinete. Simplemente, ese estado mental no se puede expresar con gritos o gemidos. La relación entre la naturaleza intrínseca de un estado mental —lo que se siente— y su expresión externa es más estrecha que eso. Sabemos que alguien que acaba de darse un golpe en un dedo del pie no está gritando porque esté viviendo una experiencia parecida a la que experimento yo cuando oigo un do medio tocado por un clarinete. Sabemos que está experimentando algo muy parecido a lo que yo experimento cuando recibo un golpe en un dedo del pie.

El argumento de la analogía con otras mentes fue el objetivo preferido de las críticas de Wittgenstein. Su principal objeción al argumento no es que sea débil, simplemente. Wittgenstein trata de mostrar que si nuestro único modo de conocer los estados mentales fuera nuestra propia experiencia, nos resultaría imposible pensar en

términos de las conciencias de otras personas. Supongamos que yo te tirase un ladrillo sobre el dedo del pie: sería como si yo pensara que ello no trae consigo ningún tipo de dolor —yo no siento ninguno—, y que eso fuera todo. Pero desde el momento en que de hecho pensamos en términos de otras mentes y sus experiencias, debemos aprender a conceptualizarlas de otro modo.

A estos efectos, el primer paso es rechazar la imagen de la mente y el cuerpo que nos ofrece el dualismo cartesiano. Y existen razones metafísicas, además de las epistemológicas, que nos animan a abandonar este dualismo. ¿Acaso el dualismo cartesiano nos ofrece una imagen razonable de cómo es el mundo, dejando a un lado lo que sabemos de él? Examinemos de nuevo al zombi. Su funcionamiento físico es idéntico al nuestro. Responde al mundo de la misma manera. Sus proyectos tienen éxito o fracasan del mismo modo que los nuestros: su salud depende de las mismas variables que la nuestra. Se pondrá a reír en el momento adecuado y llorará frente a las tragedias debidas. Puede que sea una buena compañía. ¿Qué es lo que cambia entonces debido a la falta de conciencia? O, para ponerlo al revés, ¿qué se supone que hace la conciencia en nosotros? ¿Debemos concluir que para nosotros, los no zombis, los estados mentales existen pero no hacen nada? ¿Es la conciencia algo así como la bocina en un motor: nada que sea parte de la maquinaria que lo hace funcionar? (Esta doctrina se conoce como «epifenomenalismo».) Pero si las mentes no hacen nada, ¿por qué surgieron de la evolución? ¿Por qué apostó por ellas la naturaleza? Y si los estados mentales no hacen realmente nada, ¿cómo es que los recordamos, por ejemplo?

Éste es el problema de la interacción entre la mente y el cuerpo, tal como se presenta a la luz del dualismo cartesiano.

Locke y Leibniz y el buen arbitrio de Dios

La cuestión que se plantea aquí queda magníficamente resumida en un debate que tuvo lugar entre John Locke y un contemporáneo suyo, el gran matemático y filósofo Gottfried Wilhelm Leibniz (1646-1716). Como tantos otros pensadores del siglo XVII, Locke estaba preocupado por las implicaciones de la moderna concepción científica del mundo. En particular, se interesó por la relación causal que pudiera haber entre el movimiento de ciertas partículas en el cerebro y la producción de ciertas ideas en la mente, como el color. En el pasaje que se ofrece a continuación explica cómo un bombardeo de pequeñas partículas atómicas produce efectos tales como olores, sabores, sonidos y colores.

Vamos a suponer, entonces, que los diferentes movimientos y formas, volumen y número de tales partículas, al afectar los diversos órganos de nuestros sentidos, producen en nosotros esas diferentes sensaciones que nos provocan los colores y olores de los cuerpos; que una violeta, por ejemplo, por el impulso de tales partículas insensibles de materia, de formas y volumen particulares y en diferentes grados y modificaciones de sus movimientos, hagan que las ideas del color azul y del aroma dulce de esa flor se produzcan en nuestra mente. Puesto que no es mayormente imposible de concebir que Dios haya unido tales ideas a tales movimientos con los cuales no tienen ninguna similitud, de lo que pueda serlo concebir que haya unido la idea de dolor al movimiento de un pedazo de acero que divide nuestra carne, movimiento respecto al cual esa idea de dolor no guarda ninguna semejanza.⁴

Locke compartía una concepción que ya descubrimos en Newton y en Descartes, según la cual ciertos procesos causales son relativamente inteligibles, en especial aquellos en los que una determinada cualidad, como el movimiento, pasa de una partícula a otra a causa de un impacto. Pero la relación de causalidad entre el cuerpo y la mente, de acuerdo con la cual ciertos movimientos en el cerebro producen algo enteramente distinto, a saber, las sensaciones de olor, color o dolor, resultaba completamente misteriosa. Que en un

momento dado se produzcan ciertos estados mentales resulta simplemente un hecho asombroso. Su origen sería lo que Locke llama en otro pasaje «la voluntad soberana y el buen arbitrio» de Dios, «el sabio arquitecto» que «asocia» determinadas modificaciones de la conciencia a determinados eventos físicos. Locke piensa que no tenemos una idea «clara y distinta», para decirlo en términos cartesianos, de la clase de sistemas que escogería Dios para añadirles conciencia. Que el universo esté organizado de modo que cierto tipo de sistemas, y no otros, posean alguna clase de conciencia no sería más que una cuestión fáctica. Y es también una cuestión meramente fáctica que sus conciencias cambien y adquieran determinadas propiedades en el momento en que sus cuerpos cambian y adquieren determinadas propiedades. Hay que distinguir las conexiones racionales e inteligibles, como las que encontramos en la disciplina *a priori* de las matemáticas, del hecho que ciertos «movimientos» simplemente produzcan en nosotros las sensaciones que producen. Esto último es una simple cuestión de hecho, el resultado del arbitrio divino.

En realidad, Locke no se encuentra lejos en este punto de la doctrina conocida como *ocasionalismo*, que fue abrazada por otro contemporáneo suyo, Nicolás Malebranche (1638-1715). De acuerdo con ella, no se puede decir estrictamente que los hechos físicos causen o den origen de alguna forma a los hechos mentales. Más bien representan la ocasión para que Dios mismo introduzca los hechos mentales adecuados en nuestras biografías. Hablando estrictamente, nuestros cuerpos no afectan a nuestras mentes, sólo ofrecen la ocasión para que Dios lo haga. Locke no llega propiamente a decir esto, pero podríamos pensar que existe poca diferencia entre un Dios que interviene según su arbitrio para asegurar que si un pedazo de acero divide nuestra carne, se produzca una sensación de dolor, por un lado, y un Dios que introduce directamente la sensación de dolor en el alma cada vez que un pedazo de acero divide nuestra carne, por

otro.

La doctrina de Locke disgustó profundamente a Leibniz. En el siguiente fragmento de sus *Nuevos ensayos sobre el entendimiento humano*, un pormenorizado comentario sobre Locke, éste habla por boca de Filaletes y Leibniz por la de Teófilo. Adviértase la cita directa del pasaje anterior de Locke:

Filaletes. Determinadas partículas, al impresionar nuestros órganos de una cierta manera, provocan en nosotros algunas sensaciones de color o sabor, o bien otras cualidades segundas que tienen la posibilidad de producir dichas sensaciones. «Y no es más difícil concebir que Dios pueda vincular dichas ideas (como la de calor) con movimientos, con los cuales en principio no tienen ningún parecido, que imaginar que ha vinculado la idea de dolor con el movimiento del hierro que desgarran nuestra carne, que en principio no se parece en absoluto al dolor.»

Teófilo. Tampoco hay que imaginarse que ideas como el color y el dolor sean completamente arbitrarias y no tengan ningún tipo de conexión o relación natural con sus causas: Dios no tiene la costumbre de actuar con tan poco orden y razón. Más bien diré que existe un modo de parecido, incompleto y por así decirlo *in terminis*, pero, sin embargo, expresivo, o una especie de relación de orden, como una elipse e incluso una parábola o hipérbola se parecen de alguna manera al círculo del cual son la proyección por medio de plano, puesto que entre lo que se proyecta y la proyección que de ello resulta hay una relación exacta y natural, al corresponderse cada punto de uno con cada punto de la otra según una determinada relación. Los cartesianos no han tenido esto suficiente en cuenta, y en esta ocasión os habíais apoyado en ello más de lo acostumbrado, y sin motivo alguno para hacerlo [...] Es verdad que el dolor no se parece al movimiento de un alfiler, pero se puede parecer mucho a los movimientos que dicho alfiler provoca en nuestro cuerpo, y representar dichos movimientos en nuestra alma, como no dudo que suceda.⁵

Donde Locke no ve más que el «arbitrio divino», Leibniz parece insistir en la necesidad de una conexión racional. Lo que sucede en el

alma debe mantener una relación cuasi matemática con los «movimientos» en el cerebro y en el cuerpo que son su causa.

Podemos plantear la cuestión del siguiente modo. Imaginemos a Dios creando el universo. ¿Cuánto trabajo tiene que hacer? Una doctrina atractiva sería la siguiente: debe crear los elementos físicos y las leyes de la física, y todo lo demás viene solo. Desde esta perspectiva, el Dios creador lo deja todo determinado hasta el fin de los tiempos con sólo fijar el estado físico del universo hasta el fin de los tiempos. Si hubiera querido crear un mundo en el que algo fuera distinto —por ejemplo uno donde el pinchazo de un alfiler no fuera doloroso— se hubiera visto obligado a recomponer la realidad física para que no se produjera este resultado. Hubiera tenido que escoger distintos nervios y recorridos en el cuerpo y en el cerebro. No hay ninguna variación independiente de acuerdo con la cual lo físico se mantenga inalterado y sin embargo lo mental sea distinto. Esta es la posición de Leibniz, por lo menos la que se deduce de este pasaje. (Una interpretación distinta de Leibniz le presenta convencido de que hay variaciones independientes, pero también de que Dios ha escogido, por supuesto, la mejor asociación posible entre los estados mentales y los físicos.)

Locke, en cambio, piensa que Dios tiene que hacer dos cosas distintas. Primero, determinar todo lo físico y las leyes de la física. Y segundo, decidir cómo «asociar» los estados mentales con los físicos, y establecer así las relaciones psicofísicas. Es como si el mundo tuviera dos historias distintas: una relativa a los sucesos físicos y otra relativa a los sucesos mentales, y Dios tuviera que decidir la relación entre ambas. En este caso, sería posible la variación independiente. Dios podría haber mantenido inalterada la física y optar, sin embargo, por no asociar ningún dolor al pinchazo de un alfiler.

Consideremos ahora el caso de una persona (tú mismo) y un duplicado físico de esta persona (un hermano gemelo). Si Locke está en lo cierto, entonces es posible, en principio, que tu hermano

gemelo sea un zombi o un mutante. Aunque su apariencia física sea igual que la tuya, la posibilidad de que su vida mental sea también parecida a la tuya depende de un ejercicio arbitrario de la gracia divina. Esto es particularmente obvio en la versión «ocasionalista» de esta idea: por sus propios e inescrutables motivos, puede ser que en mi caso Dios trate el golpe que he recibido en el dedo del pie como una ocasión para insertar dolor en mi biografía mental, pero puede ser que no haga lo mismo en tu caso. Por otro lado, si Leibniz está en lo cierto, tal posibilidad no existe. Si tú y tu hermano gemelo recibís un golpe igual de fuerte en el dedo del pie, y tenéis la misma reacción física, la «expresión» de tales sucesos físicos en vuestra mente debe ser también la misma, del mismo modo que las figuras proyectadas en un plano inclinado por dos formas idénticas tienen que ser las mismas.

Resulta interesante que Leibniz utilice una analogía matemática. Era mejor matemático incluso que Descartes, y entre otras cosas inventó el cálculo, pero lo importante es que para Leibniz el orden completo de la naturaleza debe ser, en último término, transparente a la razón. Las cosas no suceden de una u otra forma simplemente porque sí. Debe haber una razón para ello, sólo es cuestión de alcanzar la suficiente perspectiva de conjunto. Todo debe tener un sentido. Cuando Leibniz afirma que Dios no hace nada de forma arbitraria o carente de principios, no está expresando en realidad un determinado optimismo teológico, sino que más bien insiste en que deberíamos ser capaces de ver por qué las cosas son de una forma o de otra. Éste es su «principio de razón suficiente». Para decirlo en términos cartesianos, deberíamos ser capaces de alcanzar una idea clara y distinta del motivo por el que las cosas son como son. Deberíamos ser capaces de comprender mejor por qué la forma en que las cosas son es la forma en que deben ser. Es esta confianza en las posibilidades de la razón lo que convierte a Leibniz, igual que a Descartes, en un «racionalista».

En filosofía de la mente, el leibniziano debe negar que los zombis y los mutantes sean posibles. Si se ha fijado la biografía física, también se ha fijado la biografía mental. No hay posibilidad de variación independiente, sea actual o posible. El problema filosófico es comprender por qué eso es así. Se trata de comprender la forma en que la historia de lo físico se ajusta a la historia de lo mental.

Locke pensó que podía dejar abierta la cuestión de si es una «cosa» inmaterial (un espíritu) la que piensa dentro de nosotros, o bien si es el sistema físico mismo, puesto que Dios puede agregar pensamiento donde le apetezca. Pero deja bien claro que sólo una mente puede crear otra mente. Se requiere una concesión especial: el pensamiento no puede surgir de forma natural (o tal como lo plantea Leibniz, de una forma racionalmente explicable) a partir de la materia.

Porque las partículas no pensantes de materia, cualquiera que sea la manera en que se reúnan, no pueden, por ese motivo, recibir nada nuevo, salvo una nueva relación de posición, y es imposible que eso les comunique pensamiento y conocimiento.⁶

Es esta clase de certeza *a priori* acerca de lo que puede o no ser causa de otra cosa la que convierte en esencia a Locke, como a todos sus contemporáneos, en un racionalista, aunque tuviera mayores inquietudes acerca del poder de nuestra razón que Descartes y Leibniz.

Los demás pensadores que se han dedicado al problema de la relación entre la mente y la materia no han llegado mucho más lejos que Locke y Leibniz. También hoy existen pensadores (a veces llamados «nuevos místicos») que creen que jamás comprenderemos la relación entre la mente y la materia. Todo sigue igual a como lo dejó Locke: se trata de una relación inexplicable en términos racionales (el arbitrio de Dios). Incluso hay filósofos que defienden la existencia de algún tipo de dualismo cartesiano y que

consideran que la mente es realmente epifenoménica (no causa fenómenos físicos de ninguna clase). Su argumento es que lo físico es un «sistema cerrado». Si existe un proceso que comienza con un alfiler que se clava en la piel y termina con un quejido, entonces existe una completa cadena física que va del alfiler a la queja y que explica la queja. Luego, piensan, tiene que ser falso que nos quejemos porque sentimos dolor. Esta parcela del sentido común debe ser abandonada. Nos quejamos a causa de los procesos físicos, no a causa de un añadido mental. Estos pensadores se encuentran atascados en el mismo problema de la interacción al que se enfrentaba Locke. Trataremos esta cuestión con más detalle en el próximo capítulo.

Sin embargo, otros pensadores creen que sí se puede establecer una relación racional. A continuación voy a presentar dos grandes formas de acercarse al problema. La primera trata de ofrecer un «análisis» de lo mental, en unos términos que permitan verlo como una expresión leibniziana de lo físico. El segundo busca algún tipo de reducción o identificación científica de lo mental con lo físico.

El análisis⁷

El análisis, en el sentido que le dan los filósofos, pretende expresar las condiciones de verdad de ciertas afirmaciones misteriosas en términos algo menos misteriosos. Se puede ilustrar fácilmente lo que esto significa mediante un ejemplo familiar. Supongamos que alguien no sabe cómo interpretar este icono de la vida moderna en Occidente que es el hombre medio, con sus 2,4 hijos y sus 1,8 automóviles. ¿Qué interés puede tener un chiste como éste? Obtenemos la respuesta si mostramos las condiciones que verifican las afirmaciones implícitas en él: en este caso, que en el conjunto de las familias el número total de hijos dividido por el número de progenitores da 2,4, y que los automóviles divididos por el número de propietarios da 1,8.

Esta información queda sucintamente presentada en términos del hombre medio, que es lo que Russell llamaba una «construcción lógica» realizada a partir de un agregado de hechos. (Esto no significa que todas las afirmaciones acerca de promedios sean razonables o útiles: tal como se ha señalado, la persona media tiene un testículo y un seno.) Los filósofos hablan también de reducción de las afirmaciones de una clase a las de otra. El análisis nos permite realizar tales reducciones.

El análisis explica el significado de ciertas afirmaciones realizadas con unas palabras en términos de afirmaciones realizadas con otras palabras. Su validez como herramienta intelectual ha sido objeto de un buen número de controversias filosóficas, y su estatus ha variado a lo largo de los últimos cien años. Algunos, como Russell y G. E. Moore (1873-1958), pensaron que el análisis era el objetivo principal de la filosofía. Tales perspectivas fueron más tarde puestas en duda por el destacado pensador americano de mediados del siglo XX, W. V. Quine (n. 1908), entre otros, y su pesimismo obtuvo cierta credibilidad a partir del hecho frustrante de que muy pocos análisis filosóficos resultaban satisfactorios. Su situación actual se puede interpretar como un prudente renacimiento. Sin embargo, nuestro propósito actual nos permite dejar a un lado estas cuestiones metodológicas. Lo importante es que si logramos analizar las asociaciones mentales en términos físicos, se renueva el sueño leibniziano de una explicación racional y *a priori* de cómo lo físico da origen a lo mental.

Tomemos el dolor como ejemplo de un determinado estado mental. Supongamos ahora que intentamos analizar qué significa para una persona sentir dolor. Identificamos el dolor ante todo a partir de aquello que el dolor nos mueve a hacer (lo cual es también su razón de ser, en términos evolucionistas). El dolor nos mueve a hacer un cierto número de cosas. Reclama atención, hace que dejemos inmóviles algunas partes del cuerpo, nos distrae de otras

cosas y, por supuesto, es desagradable. Supongamos que somos capaces de resumir estas consecuencias en términos de tendencias o disposiciones de conducta. Cabe sugerir entonces que sentir dolor no es más que tener aquellas disposiciones de conducta. Éste sería el análisis de lo que significa sentir dolor, es decir, de sus condiciones de verdad. Esta conclusión sería el resultado de un ejercicio *a priori* de la razón, la reflexión acerca de lo que realmente significan las afirmaciones sobre esta clase de estado mental. Desde esta perspectiva, el misterio de la conciencia desaparece. Dado que tú y tu hermano gemelo compartís ciertas disposiciones (tendéis de forma observable hacia el mismo comportamiento), compartís también vuestras sensaciones, pues en eso consisten precisamente las sensaciones. Esta doctrina se llama «conductismo lógico». Estoy convencido de que posee cierta verdad, aunque también tiene algunos problemas. Una posible objeción se basa en un hecho que a todos nos resulta familiar: las personas pueden compartir la misma sensación y sin embargo reaccionar de forma algo distinta. Puede ser que un día nos demos un golpe en el dedo del pie y montemos por ello un terrible espectáculo, y quizás otro día nos pase lo mismo, sintamos el mismo dolor y, sin embargo, saquemos una valiente sonrisa y sigamos adelante. El comportamiento no es una guía infalible hacia las sensaciones, los pensamientos o los sentimientos. (Éste es el sentido de aquel chiste sobre dos conductistas que estaban en la cama: «Te lo has pasado muy bien, ¿cómo me lo he pasado yo?».) O sea, que como mínimo habrá que introducir alguna complicación. Tal vez podamos salvar el análisis en términos de disposiciones de conducta señalando que sigue siendo cierto que estamos, en algún sentido, dispuestos a realizar demostraciones de dolor más expresivas, aunque las estamos reprimiendo por una u otra razón. Resulta prácticamente imposible reprimir por completo las tendencias de conducta relacionadas con el dolor, y una tercera persona no tiene problemas para distinguir, por ejemplo, un niño que

no se ha hecho daño de otro que sí pero que se está haciendo el valiente. Esta clase de disposiciones de conducta parecen ser un rasgo esencial del dolor. Pero existen ciertos casos de lesiones cerebrales que arrojan dudas incluso sobre este punto, ya que aquellos que las sufren afirman de forma aparentemente sincera que sienten todavía algún dolor, pero que ya no les importa. Hay que llamar la atención, sin embargo, sobre lo difícil que resulta darle algún sentido a esta afirmación. Si pensamos en algún ejemplo de un fuerte dolor —tocar una placa de cocina o dar una patada a una pared— resulta difícil imaginar aquel estado mental en concreto sin imaginarlo como algo increíblemente desagradable, como también resulta difícil imaginarlo sin su tendencia a provocar manifestaciones típicas de dolor.

Los pensadores contemporáneos tienden a no poner demasiada fe en esta clase de conductismo. Prefieren una doctrina ligeramente más elaborada que se conoce como «funcionalismo», y que presta también una especial atención a la función de los estados mentales. Sin embargo, concibe esta función de un modo algo más flexible. Parte de la idea de que habría una red de relaciones físicas: no sólo disposiciones de conducta, sino causas típicas e incluso efectos sobre otros estados mentales, en el supuesto de que éstos también se puedan traducir en una serie de disposiciones físicas. La idea, sin embargo, es muy parecida.

El dolor es un estado mental que se presta de un modo especial a los proyectos de análisis, ya que tiene una expresión razonablemente característica y natural en la conducta. Otros estados con el mismo tipo de expresión natural son las emociones (la tristeza, el miedo, el enfado o la alegría tienen también manifestaciones típicas en la conducta). En cambio, otros estados mentales sólo mantienen una relación indirecta con la conducta: piénsese en el sabor del café, por ejemplo. Saborear un café constituye una experiencia característica. Hay algo que para nosotros corresponde a saborear un café (no en el

caso de los zombis). Pero no nos incita típicamente a hacer demasiadas cosas. A los pensadores contemporáneos les gusta decir que hay *qualia* o sensaciones en bruto asociadas con el sabor del café. Y los defensores de estos *qualia* acostumbran a ver con malos ojos los proyectos de reducirlos a disposiciones de conducta. A estos efectos, vuelven a la posición de Locke. Se daría el caso de que estos *qualia* serían algo sobreañadido a ciertos hechos físicos —al menos en mi caso, si no en el tuyo— aunque podría haber sido de otro modo. Pero en ese caso amenaza de nuevo el escepticismo acerca de si vosotros sois o no sois zombis o mutantes.

Un modelo científico

Es importante señalar una distinción que aparece a menudo en el debate contemporáneo. Hasta ahora hemos presentado a Leibniz como un enemigo del elemento de facticidad en bruto que se daba en Locke, en nombre de una relación racional y cuasi matemática entre la mente y el cuerpo. Se podría sugerir que hay una ruta intermedia: opuesta a la facticidad, pero no hasta el punto de proponer una relación matemática o racionalmente transparente. El planteamiento habitual es decir que tal vez existe una identidad metafísica entre los hechos mentales y los físicos, pero no es necesario que se pueda conocer *a priori*.

Es habitual proponer la siguiente analogía. La física clásica identifica la temperatura de un gas con la energía cinética media de las moléculas que lo componen. De modo que para crear los gases calientes Dios no tiene que hacer más que una cosa: definir el gas y la energía cinética media de sus moléculas, con lo cual fija también la temperatura. No hay posibilidad de variación independiente. No puede haber gases zombis o mutantes, en los que la energía de las moléculas o bien no resulte en ninguna temperatura en absoluto, o bien resulte en temperaturas diferentes de las que se asocian con la

misma energía en otros gases.

Por otro lado, no ha sido sólo la razón, el pensamiento o las matemáticas lo que ha permitido a los científicos equiparar la temperatura con la energía cinética media. El paso no se dio *a priori*, mediante un análisis autoevidente del significado de temperatura, sino que fue preciso experimentar, observar y realizar consideraciones teóricas generales. El resultado no se obtuvo puramente *a priori*, sino en gran medida *a posteriori*. No es posible adivinar la relación por adelantado, contando sólo con las matemáticas o con «ideas claras y distintas», tal como sabemos por ejemplo que una circunferencia situada sobre un plano inclinado proyecta una elipse.

Por lo general, cuando en el ámbito de la ciencia se identifica un término o propiedad teórica, como por ejemplo la temperatura, con otro (en este caso la energía cinética media de las moléculas constituyentes), establecemos la relación a partir de una serie de principios-puente que forman parte del bagaje teórico de la ciencia en cuestión. Así, por ejemplo, la identificación actual de los genes con fragmentos de ADN se debe a que en biología clásica los genes se definen a partir de su función en la transmisión hereditaria de ciertas características, y en la actualidad la biología molecular establece que tal función corresponde a los fragmentos de ADN. Adviértase que en este caso el análisis no se halla completamente ausente. Es preciso saber lo que se supone que deben hacer los genes antes de que sea posible realizar la identificación. Pero el gran descubrimiento consiste en el hallazgo científico, contingente, del elemento específico que cumple la función prevista en la definición.

Si abordáramos el problema de la mente y el cerebro a partir del tipo de reducciones científicas que acabamos de describir, descubriríamos que las personas que comparten determinados estados mentales manifiestan también ciertos estados físicos característicos. Así, por ejemplo, descubriríamos que todas las

personas que sienten dolor, y sólo ellas, comparten cierto estado cerebral que a menudo se describe vagamente diciendo que sus «fibras C se están activando». Y por lo tanto se podría proponer que el estado de dolor es eso mismo, igual que los genes son determinados fragmentos de ADN. Una vez más, lo mental quedaría enteramente reducido a lo físico. Esto es lo que se conoce como «teoría psicofísica de la identidad».

Los que se oponen a esta teoría dicen a veces que sólo se puede creer en ella al precio de fingir una permanente anestesia. El motivo de su protesta es que según ellos se habría ignorado todo lo específicamente mental. La forma adecuada de responder a esta objeción es preguntar al oponente qué es exactamente lo que piensa que se ha ignorado y contemplar cómo se debate en los dilemas del dualismo. Pero existen otras dificultades relacionadas con esta clase de teorías de la identidad psicofísica. Una de ellas es que cuando se trata de estados mentales, la última palabra la tiene la propia conciencia, en el sentido que veremos a continuación. Desde la perspectiva del sujeto, cualquier cosa que produzca la misma sensación que el dolor es dolor —no importa si son las fibras C o cualquier otra cosa—. Si se le realizara un minitrasplante a una persona, por ejemplo, de modo que las fibras C orgánicas fueran reemplazadas por algún dispositivo de silicio y se diera el caso de que el silicio produjera los mismos efectos, seguiría siendo dolor. Nuestro conocimiento del dolor no depende de si tenemos fibras C en nuestro interior o cualquier otra clase de arquitectura biológica. La autoridad la tiene el que habla en primera persona. Del mismo modo, no nos sentiríamos muy cómodos si tuviéramos que decidir si la gamba siente o no siente dolor, por poner un ejemplo escasamente intuitivo, a partir del dato exclusivo de si tiene o no fibras C. De modo que la identificación no resulta tan evidente como en otros casos que ha estudiado la ciencia (esto último se podría discutir).

Nos sentiríamos razonablemente satisfechos si consiguiéramos

comprender la relación entre los estados mentales y los estados del cerebro o del cuerpo con la misma claridad que comprendemos la relación entre la temperatura de los gases y su energía cinética media. Tal vez no nos preocuparía demasiado la cuestión de si el resultado se ha logrado gracias al «pensamiento puro» o más bien a través de la experimentación. De modo que podemos estar de acuerdo con la objeción que Leibniz le hizo a Locke, sin necesidad de compartir enteramente el racionalismo de aquél. Sin embargo, cuando tratamos de pensar seriamente la relación entre el cerebro y el cuerpo, por un lado, y la mente, por el otro, a menudo parece que lo que nos impide salir adelante es nuestra forma de pensar, más que una mera ignorancia científica. Recientemente, un buen número de científicos ha dirigido su atención a la conciencia y se han identificado un cierto número de estados cerebrales implicados en el funcionamiento normal de la conciencia. Por ejemplo, se cree que es fundamental la presencia de ondas electromagnéticas de una frecuencia particularmente baja en el cerebro. Pero no está claro si esta clase de evidencia puede servir para resolver el problema, es decir, para que podamos ponernos al lado de Leibniz en su enfrentamiento con Locke. Desde el punto de vista de este último, todo lo que han descubierto los científicos es que cuando el cerebro se encuentra en un estado específico, se dan síntomas de conciencia. Pero acaso eso no sea más que una asociación puramente fáctica entre la conciencia y un cierto estado. No lo convierte en una combinación inteligible. Por otro lado, eso nos permite descartar las opciones zombi y mutante, pues de otro modo los científicos no habrían podido establecer ninguna correlación, excepto quizá para este o aquel caso. Sin embargo, de acuerdo con los nuevos misterios, ni la ciencia ni la filosofía conseguirán jamás llegar a un punto en el que las cosas estén más claras. Nunca podremos ponernos por entero al lado de Leibniz en su enfrentamiento con Locke.

Espectros invertidos: lenguajes privados

En muchos casos parece que la cuestión de los colores abre la puerta, cuando menos, a los mutantes (personas idénticas a las demás desde el punto de vista físico y que sin embargo perciben los colores de forma harto distinta). Incluso podría haber mutantes que tuvieran completamente invertidos los respectivos espectros de colores, de modo que la experiencia de la luz que obtiene uno al extremo rojo del espectro fuera la misma que obtiene el otro al extremo azul. Y no habría forma de explicarles que esto es así.

El dualismo cartesiano abre la posibilidad de que haya zombis y mutantes. Pero quizá también abre otra posibilidad todavía más escalofriante. Desde una perspectiva dualista, podemos sentirnos seguros al menos de conocer nuestra propia experiencia. Tal vez la mente de los demás resulte algo incierta, pero la nuestra la conocemos bien. Pero ¿es eso cierto? No pensemos ahora en la mente de los demás, sino en nuestra propia experiencia pasada. ¿Estamos seguros de que hoy vemos el mundo del mismo color que ayer? ¿Estamos realmente seguros de que ayer poseía algún color, en otras palabras, de que tuvimos la experiencia consciente que recordamos haber tenido?

Cuando nos planteamos esas preguntas, lo que hacemos es aplicar las opciones zombi y mutante a nuestro propio pasado. No hay duda de que a primera vista tales hipótesis resultan todavía más peregrinas y absurdas que cuando las aplicábamos a las otras mentes. Y nos sentimos inclinados a replicar que por supuesto ayer veíamos más o menos los mismos colores que hoy. Seguramente nos daríamos cuenta si al levantarnos por la mañana el cielo tuviera el mismo aspecto que tenía la hierba ayer, y viceversa.

Por supuesto, estoy de acuerdo en que nos daríamos cuenta de la diferencia. Pero ¿hay algo que garantice esta seguridad, si admitimos el dualismo cartesiano? Depende de lo que pensemos acerca de la

memoria y de los estados mentales. ¿Por qué habríamos de estar seguros de que los estados mentales —recordemos que se conciben como algo enteramente distinto de todo lo físico— dejan huellas fiables en la memoria? Puedo comprobar si mi recuerdo del mundo físico es lo bastante fiable. Recuerdo haber puesto el coche en el garaje, y he aquí que cuando bajo a mirar, allí lo encuentro. Recuerdo el camino para ir hasta la cocina, y he aquí que llego hasta ella apenas sin esfuerzo y sin cometer ningún error. Pero ¿de qué modo podría comprobar si mi recuerdo del mundo mental es correcto? En términos de Locke, ¿acaso el «buen arbitrio» divino no puede introducir en mí ciertas modificaciones mentales hoy, juntamente con el falso recuerdo de que ayer conocí otras parecidas? Wittgenstein dijo:

Líbrate siempre del objeto privado asumiendo: está cambiando continuamente; pero tú no lo notas, porque tu memoria te engaña continuamente.⁸

Éste es el corazón del argumento de Wittgenstein contra el lenguaje privado, uno de los argumentos más celebrados de la filosofía del siglo XX, que aparece en sus *Investigaciones filosóficas* (publicadas a título póstumo en 1953). Wittgenstein pretendía hacernos ver que si nuestra vida mental estuviera divorciada del mundo físico, tal como propone el dualismo cartesiano, no habría nada relevante que pensar acerca de la naturaleza de la propia vida mental en el pasado (o en el futuro). Sería, por así decirlo, un terreno demasiado resbaladizo o misterioso como para ser objeto de nuestros propios recuerdos o intenciones.

Las opciones zombi y mutante, aplicadas a nuestro propio pasado, resultan ciertamente descorazonadoras. Pero en realidad sólo deberían hacernos perder la confianza en la propuesta dualista. La pregunta sigue siendo: ¿podemos dar marcha atrás desde la posición de Locke hacia alguna versión de la posición de Leibniz? Recordemos

que Leibniz pretende que existe una relación «racional» entre lo físico y lo mental, de modo que el estado mental de ver un color sería una determinada expresión racional de lo que sucede a nivel físico, y no un añadido accidental. ¿Cómo podemos aplicar esto al caso de los colores? La idea de Leibniz es que si tanto yo como mi hermano gemelo (que en este caso podría ser yo mismo tal como era ayer) funcionamos de la misma manera a nivel físico, entonces no es posible que nuestras vidas a nivel mental sean distintas. ¿Cómo podríamos desarrollar esta propuesta? Veamos una posible respuesta.

Tenemos una comprensión razonablemente buena de muchos de los cambios físicos que tienen que ver con la percepción del color. Esta es el resultado de la estimulación de unos conos que se encuentran en la parte central de la retina. La mejor teoría que tenemos en la actualidad sugiere que existen tres tipos distintos de cono, L, M y S (grande, mediano y pequeño).⁹ Los conos L tienden a «pinchar» o enviar mensajes a través del nervio óptico cuando reciben ondas lumínicas de longitudes más largas; los conos M se excitan más bien cuando se trata de luz de longitudes medias, y los conos S cuando se trata de luz de longitudes más cortas. El color que percibimos depende, por lo tanto, en primer lugar, de la comparación de los niveles de excitación de estos tres tipos de cono. Así, por ejemplo, si S está mucho más excitado que L, se trata del código del azul, el color que se encuentra en el extremo de longitud corta del espectro. Si L está mucho más excitado que S, es el código del amarillo. Si L está más excitado que M, se trata del rojo, y si M está más excitado que L, del verde. Es algo así como si los canales fueran «adversarios» y el resultado dependiera de cuál de ellos vence.

Los colores poseen muchas propiedades interesantes. Veamos algunas de ellas: no es posible ver una superficie azul-amarillenta. Tampoco una que sea verde-rojiza. Por otro lado, sí es posible una superficie verde-azulada o rojo-amarillenta (naranja). No es posible un marrón brillante. Tampoco un gris brillante (resulta difícil imaginar

una llama gris o marrón). El amarillo es un color más claro que el violeta. Es posible que exista una piedra preciosa transparente de color rojo, azul o verde, pero no es posible que exista una piedra preciosa transparente de color blanco —lo más parecido sería un blanco lechoso, como un ópalo—. Es posible que la luz sea blanca, pero no que sea negra.

Podría parecer que todo esto no es más que un conjunto de hechos en bruto acerca del reino cartesiano de la mente, donde se supone que residen los colores. Pero también podemos comenzar a verlo como la expresión de una serie de hechos físicos. No podemos ver una superficie de color azul-amarillento porque el amarillo y el azul proceden de situaciones lógicamente opuestas: el amarillo corresponde a $L > S$, y el azul a $S > L$. Sucede algo parecido con el rojo y el verde. No es posible un marrón brillante porque el marrón es amarillo oscurecido. Una superficie se ve marrón cuando el código correspondería al amarillo, si no fuera porque en líneas generales hay un bajo nivel de energía en comparación con el de otras fuentes cercanas de luz. Sucede algo parecido con el gris, que es blanco oscurecido. El amarillo es más claro que el violeta porque la luz amarilla ($L > S$) se halla también más cerca de la frecuencia en la que nuestro sistema visual ofrece su máxima respuesta. Comparativamente, tanto el rojo como el azul, a ambos extremos del espectro visual, se acercan al punto donde ya no ofrecemos ninguna respuesta, lo oscuro. No es posible un blanco transparente porque sólo podemos ver algo blanco cuando dispersa la luz.

Por supuesto, todo esto no es más que una aproximación superficial a nuestro conocimiento científico de los colores. Pero al menos nos permite vislumbrar el «sentido» que tienen las cosas. Si dispusiéramos de suficientes hechos de este tipo, quizá nos sentiríamos menos fascinados por la hipótesis del espectro invertido. Consideremos primero el caso más simple de visión monocromática (en blanco y negro). Supongamos que se plantea la posibilidad de que

alguien sea un duplicado físico de mí mismo, pero que vea oscuridad allí donde yo veo luz, y viceversa. ¿Es eso posible? Nuestra primera respuesta quizá sería que sí. Tal vez nos imaginaríamos que ve el mundo tal como se ve en un negativo fotográfico. Pero eso, en realidad, no tiene sentido. Veo mejor a través de una pieza de cristal gris si la hago más clara; si la hago más oscura, veo peor. Como mi hermano gemelo es un duplicado físico mío, esto también es cierto en su caso. Para él, sin embargo, cuando hacemos más claro el cristal, «parece» como si le hubiéramos puesto hollín, puesto que subjetivamente se ha vuelto más oscuro. Y cuando ponemos hollín «parece» como si se volviera más claro. Pero aun en ese caso debemos imaginar que para él, cuando el plato de cristal se vuelve más oscuro, ve mejor a través de él, y cuando se vuelve más claro, ve cada vez peor. Y eso simplemente no parece tener sentido. No constituye una hipótesis coherente.

Consideremos ahora el caso de alguien que sea físicamente idéntico a mí, pero que supuestamente vea amarillo donde yo veo azul, y viceversa. En este caso no resulta tan fácil imaginárselo. Tiene que reaccionar de la misma forma que yo, de modo que no puede ir diciendo que el amarillo es un color oscuro, por ejemplo. Tal diferencia en las reacciones y en el comportamiento sería una diferencia física, de modo que tenemos que preguntarnos cómo puede ser que vea el azul brillante y el amarillo oscuro. Si realmente ve oscuro el amarillo, tal como yo veo el azul, ¿cómo ve el marrón? ¿Cómo ve el naranja? El marrón es amarillo oscurecido, pero para él el amarillo ya es oscuro. Así, resulta difícil imaginar cómo sus juicios acerca del mundo físico podrían coincidir con los míos si partimos de esta completa disparidad de experiencias mentales.

En resumen, tal posibilidad resulta mucho menos clara, y puede que nos sintamos inclinados a negar que sea en absoluto una posibilidad. Con ello estaríamos sentando las bases para una concepción de la mente que eliminaría la distancia entre el mundo

físico y el mental, esto es, entre el activo y sensible sistema visual de nuestro cerebro y los «subjetivos» *qualia* aparentemente sobreañadidos de la experiencia de color. Un diseño de este tipo constituiría una reivindicación de la postura de Leibniz. La experiencia subjetiva del color no dejaría de ser un mero añadido estrambótico, sino que se convertiría en la expresión inevitable, racionalmente explicable, del tipo de funcionamiento físico propio de las criaturas que somos nosotros. Si se puede hacer lo mismo con todos los elementos de nuestra conciencia, el problema está resuelto.

El pensamiento

Vamos a tratar ahora un aspecto de la conciencia ligeramente distinto. En este capítulo nos hemos concentrado en las sensaciones y en los *qualia*. Pero nuestra conciencia se compone también, en gran medida, de pensamientos. Los pensamientos son algo extraño. Tienen capacidad «representativa»: típicamente, un pensamiento representa el mundo de una u otra forma. En cambio, parece que una sensación simplemente está allí. Según parece, no remite a nada más allá de sí misma, a ningún hecho real o imaginario. (Algunos pensadores no están de acuerdo con esto. Piensan, por ejemplo, que la sensación de dolor es la percepción de una herida en el cuerpo, y que esta percepción representa el cuerpo como un cuerpo herido, del mismo modo que el pensamiento de que mañana es viernes representa el día de mañana como si fuera viernes. Dejo al lector que considere lo plausible que pueda ser esta propuesta.) La naturaleza representativa de los pensamientos, también llamada «intencionalidad» o «direccionalidad», es ella misma altamente desconcertante. Si consideramos que los pensamientos son un cierto tipo de «cosas» que se hallan presentes en nuestra conciencia, la pregunta es cómo es posible que una «cosa» señale en y por sí misma hacia otra cosa (un hecho o un estado de cosas). Ciertamente,

una señal de tráfico puede referirse a un pueblo, por ejemplo. Pero eso parece que depende de la forma en que es interpretado. Una señal de tráfico no representa en y por sí misma el camino hacia el pueblo. Debemos aprender a interpretarla del modo adecuado. Podemos imaginarnos una cultura en la que el mismo objeto físico que para nosotros es una señal de tráfico tuviera una función hartamente distinta: un panel para colgar anuncios, un tótem o una pieza de arte abstracto. Lo mismo sucede con los animales: cuando señalamos algo, es típico que los perros sólo presten atención al dedo que señala, para mayor irritación de su dueño. En cambio, resulta incoherente imaginar una criatura con los mismos pensamientos que nosotros que no haya aprendido a interpretar esos pensamientos como lo hacemos nosotros. El pensamiento consiste en esta «interpretación».

Es probable que la mejor forma de responder a esto sea negar que los pensamientos sean cosas. El error de suponer que detrás de cada nombre hay una «cosa» recibe a veces el nombre de *reificación*. Es frecuente que los pensadores se acusen unos a otros de caer en falsas reificaciones. Son las personas las que piensan, y ello no depende de ninguna especie de mancha que haya que buscar en el cerebro o en la mente. Esto es cierto incluso si se piensa en esta mancha como si fuera una breve oración escrita en el cerebro. Pensar consiste en interpretar el mundo de una forma o de otra y, por lo tanto, esta interpretación depende más de nuestras inclinaciones que de las cosas que pueda haber en nuestro interior.

Tal vez el hecho de que seamos capaces de pensar en estados de cosas ausentes —estados de cosas distantes, pasados y futuros— no sea más extraño que el hecho de que seamos capaces de percibir algo acerca del mundo y, sin embargo, es una hazaña que nos distingue de los demás animales. Es presumible que los animales perciben el mundo, pero nos resulta incómodo pensar que puedan representarse estados de cosas distantes, pasados y futuros; y, sin

embargo, nosotros somos ciertamente capaces de hacerlo.

En la actualidad, la forma más popular de abordar este problema se basa en nuestro modo de atribuir pensamientos a una persona en su estado normal. Lo que nos permite interpretar qué piensa acerca del día anterior, o acerca de la predicción meteorológica para el fin de semana, tiene que ser algo relacionado con su comportamiento. Los pensamientos se expresan tanto en el comportamiento lingüístico como en el no lingüístico, y tal vez podamos confiar en encontrar algún tipo de reducción: «X piensa que p » si y sólo si los planes, los deseos o el comportamiento de X resultan en alguna medida consistentes con el hecho de que el mundo sea tal que p . La clave sería dar contenido al «en alguna medida consistentes». Es justo decir que nadie lo ha conseguido de forma satisfactoria, aunque se han hecho algunas sugerencias acerca de la mejor forma de abordar el problema. Decimos que un sistema inteligente, como por ejemplo un misil guiado, piensa que hay un avión a la distancia de una milla y doscientos pies, si el sistema apunta hacia un lugar en el que resulta apropiado que haya un avión —en el supuesto de que su objetivo (o función) sea abatir aviones—. De modo parecido, podríamos decir que una persona piensa que el tiempo será bueno el fin de semana si su comportamiento es apropiado, dados sus objetivos (o funciones), al hecho de que el tiempo sea bueno el fin de semana. La dificultad consiste en dar algún contenido a esta idea sin remitirse de algún modo a otros estados mentales del sujeto, y esto es lo que nadie sabe cómo hacer.

Dejemos a un lado por ahora la cuestión del pensamiento. En su lugar, en los próximos dos capítulos examinaré dos nuevos aspectos de nuestra concepción del mundo que alimentan el dualismo cartesiano. El primero consiste en una serie de reflexiones acerca de nuestra propia libertad. El segundo consiste en una serie de reflexiones acerca de nuestra propia identidad.

CAPÍTULO

3

La libertad

En fin, si todos los movimientos se encadenan y el nuevo nace siempre del anterior, según un orden cierto, si los átomos no hacen, declinando, un principio de moción que rompa las leyes del hado, para que una causa no siga a otra causa hasta el infinito, ¿de dónde ha venido a la tierra esta libertad que *gozan* los seres vivientes?

Lucrecio, *De la naturaleza*.¹

O bien, en un tono ligeramente menos elevado:

Un hombre joven dijo: ¡Es el colmo!
He descubierto que soy
una criatura que avanza
siguiendo una ruta predestinada:
no un autobús, sino un tranvía.

En el último capítulo hemos reflexionado acerca de lo que produce el cerebro: elementos de la conciencia como los pensamientos, las sensaciones o los *qualia*. Sin embargo, también tenemos presentes otras cosas cuando pensamos acerca de nosotros mismos. No nos limitamos a registrar el mundo, tal como creemos que es. Actuamos

en él. Nos planteamos posibles alternativas. Deliberamos y hacemos cosas. Controlamos la situación. ¿Qué deberíamos pensar acerca de eso?

Los lazos del destino

Habitualmente nos consideramos libres. Vivimos nuestras vidas en un horizonte abierto de posibilidades. Deliberamos acerca de cuáles perseguir y, una vez que hemos deliberado, tomamos una decisión. Este año pasé las vacaciones en las montañas, pero podría haber ido a la costa. Fue una decisión mía. No podría haber ido a la Luna, porque no era factible.

Parece que somos conscientes de nuestra libertad. La conciencia de la propia libertad parece estar estrechamente asociada a cualquier clase de conciencia. Cuando en el capítulo anterior reflexionábamos acerca de los zombis, es probable que nos imagináramos algo así como unos torpes y estúpidos robots, movidos por algún programa prefijado de comportamiento. Pero nosotros no somos así, ¿verdad que no?

A veces nos sentimos orgullosos de nuestra libertad: no somos esclavos del instinto y del deseo. Podemos enfrentarnos a nuestras obsesiones o adicciones y tratar de dominarlas. Cuando lo conseguimos, merecemos la admiración de los demás. Si fracasamos, merecemos un castigo, y a veces lo recibimos. La libertad trae consigo la responsabilidad, y aquellos que abusan de ella merecen ser condenados y castigados. Pero nadie merece ser castigado por no hacer algo que no podía hacer. Sería del todo injusto castigarme por no haber ido a la Luna, o castigar a un hombre encarcelado por no haber acudido a una cita fuera de la cárcel, por ejemplo. En estos casos los obstáculos están fuera del control del agente, lo cual significa que está libre de culpa.

Así, tanto nuestros juicios morales como nuestra forma habitual de

pensar parecen presuponer que en ciertas ocasiones depende de nosotros que cometamos o no una mala acción.

¿Pero acaso no es posible que nos engañemos acerca de nuestra libertad? ¿Realmente podríamos haber actuado alguna vez de forma distinta a como lo hicimos?

Tanto Lucrecio como el joven que aparece al comienzo del capítulo parecen defender el siguiente argumento:

El pasado controla el presente y el futuro.

No podemos controlar el pasado.

Tampoco podemos controlar la forma en que el pasado controla el presente y el futuro.

Luego, no podemos controlar el presente y el futuro.

En realidad no controlamos nada en absoluto: ni el pasado, ni el presente ni el futuro.

La primera premisa de este argumento es una versión simplificada de la doctrina que se conoce como «determinismo», que podemos resumir diciendo que todo hecho es el resultado de una serie de causas precedentes. El estado del mundo en cada momento es el resultado de su estado inmediatamente anterior y evoluciona a partir de este estado precedente siguiendo una serie de leyes inmutables de la naturaleza. La segunda premisa parece ser verdadera. La tercera nos recuerda que no tenemos ningún control sobre las leyes de la naturaleza, sobre la forma en que los hechos dependen unos de otros. Y ciertamente parece que la conclusión se sigue de las premisas.

Las personas que aceptan este argumento reciben el nombre de «deterministas» radicales o «incompatibilistas», ya que creen que la libertad y el determinismo son incompatibles.

Si queremos restaurar la libertad humana, acaso debemos rechazar primero el determinismo. Tenemos razones para ser optimistas, ya que la ciencia de la naturaleza más destacada del

momento, la física cuántica, postula la existencia de hechos incausados, al menos según la interpretación más extendida. En el mundo de los cuantos, existen hechos microfísicos que «pasan porque sí». De acuerdo con esta interpretación, puede darse el caso de que un sistema se encuentre exactamente en el mismo estado que otro —no hay «variables ocultas»— y, sin embargo, que en uno de los sistemas se produzca un determinado suceso cuántico y en el otro no. Esta clase de sucesos no tienen ninguna causa: simplemente ocurren, o no ocurren. La física cuántica les asigna una probabilidad, pero no puede determinar, a partir de las condiciones existentes en un determinado momento, si tal suceso se producirá o no en un futuro inmediato.

Sin embargo, esto no es exactamente lo que buscábamos, ya que introduce un elemento de azar en las cosas, pero no un elemento de control o de responsabilidad. Veremos mejor a qué me estoy refiriendo si pensamos en el sistema neurofisiológico formado por nuestro cerebro y nuestro cuerpo. Todo sucede de acuerdo con una causa. Si a un nivel microfísico se producen a veces cambios al azar, difícilmente se nos puede considerar responsables de los cambios que se hayan producido a consecuencia de ellos. No podemos controlar los saltos de los electrones. Siendo algo genuinamente indeterminado, no hay nada que pueda controlarlos. Mala suerte si uno salta en la mala dirección: sería como si nuestras buenas intenciones se hubieran frustrado por culpa de un accidente externo ajeno a nuestro control. Introducir el accidente en nuestro cerebro no nos convierte en responsables.

El indeterminismo físico convierte la responsabilidad y la ética de la culpa en algo todavía más escurridizo. Es lo que se conoce como «el dilema del determinismo». Si el determinismo se sostiene, perdemos nuestra libertad y nuestra responsabilidad. Si no se sostiene, ya que algunos sucesos «pasan porque sí», perdemos igualmente nuestra libertad y nuestra responsabilidad. El azar es tan

ciego como la necesidad.

Higueras y cascadas

En el Evangelio de san Marcos (11:12-14, 20-21) se puede leer una historia bastante extraña:

Al día siguiente, cuando salieron de Betania, sintió hambre. Y viendo de lejos una higuera con hojas, fue a ver si encontraba algo en ella; acercándose a ella, no encontró más que hojas; es que no era tiempo de higos. Entonces dijo a la higuera: «¡Que nunca jamás coma nadie fruto de ti!». Y sus discípulos oyeron esto.

Al pasar muy de mañana, vieron la higuera, que estaba seca hasta la raíz. Pedro, recordándolo, le dice: «¡Rabbí, mira!, la higuera que maldijiste está seca».

Ignoremos los graves problemas sociales, económicos y ecológicos que plantea la historia y concentrémonos en la aparente injusticia cometida con la higuera. Es cierto que Jesús no maldijo a la higuera por no dar, por ejemplo, manzanas o ciruelas. Lo que quería eran higos. Y las higueras a veces dan higos. Y sin embargo sigue pareciendo injusto con la higuera. Es como si Jesús argumentara: «A veces das higos, luego podrías hacerlo ahora». Frente a lo cual resulta perfectamente adecuado defender a la higuera diciendo que da higos en verano, pero que entonces era invierno, o que en cualquier caso «no era tiempo de higos». Se requieren ciertas condiciones para que una higuera dé higos: ni siquiera la mejor de las higueras puede hacerlo fuera de temporada, del mismo modo que no puede dar ciruelas.

Es posible que la higuera no estuviera al corriente de esto. Quizá si hubiera sido una higuera pensante, se habría sentido mal porque no conocería las condiciones exactas que se requieren para dar higos: quizá sólo recordaría que a veces lo hace y se sentiría mal por no

hacerlo en esta ocasión. Pero eso no es más que ignorancia. Sería irracional que la higuera se sintiera mal por no dar higos en invierno: no era el momento adecuado, nada más.

Podríamos pensar igual que nuestra higuera imaginaria: sólo sé que soy libre. Estoy aquí, puedo levantar el brazo o no levantarlo a mi antojo. Supongamos que lo hago —así— y que siento cómo controlo el desarrollo de los acontecimientos. Mi conciencia me revela mi propia libertad.

Y sin embargo, aquí tenemos al filósofo alemán Schopenhauer(1788-1860):

Imaginemos a un hombre que, en medio de la calle, se dijera a sí mismo: «Son las seis de la tarde, la jornada laboral ha terminado. Ahora puedo ir a dar un paseo, o puedo irme al club; también puedo subirme a la torre para contemplar la puesta de sol; puedo ir al teatro; puedo visitar a tal o cual amigo; puedo incluso salir de la ciudad, hacia el ancho mundo, y no volver jamás. Todo esto depende únicamente de mí, mi libertad es completa. Y sin embargo no voy a hacer ninguna de estas cosas ahora, sino que por mi libre decisión iré a casa junto a mi esposa». Sería lo mismo que si el agua se dijera a sí misma: «Puedo formar olas de gran altura (¡claro que sí!, durante una tormenta en el mar), puedo lanzarme montaña abajo (¡sí!, en el cauce del río), puedo precipitarme hecha espumas y borbotones (¡por supuesto!, en la cascada), puedo levantarme como un río en el aire (¡sí!, en la fuente), y puedo, finalmente, hervir y desaparecer (¡sí!, a cierta temperatura); pero no voy a hacer ninguna de estas cosas ahora, y voy a permanecer voluntariamente quieta y clara como un espejo en este estanque».²

En esta parábola, el agua no es consciente de las estructuras causales necesarias para que hierva, forme olas y demás. Sólo recuerda que a veces hace esas cosas. Por lo tanto, piensa, soy capaz de hacerlas. De modo que atribuye el hecho de estar en reposo a una decisión personal suya. Pero en esto se equivoca: si «intentara» hervir cuando la temperatura no fuera la adecuada, o «intentara»

formar olas en ausencia de viento, pronto descubriría que esas cosas no dependen de su decisión personal. En el mismo sentido, Wittgenstein imaginaba una hoja empujada por el viento de otoño que se dijera a sí misma: «Ahora iré en esta dirección, ahora en esta otra».

Schopenhauer niega que nuestro propio autoconocimiento, nuestra autoconciencia, revele una libertad real. Podemos interpretar que su crítica se dirige contra el siguiente argumento:

No soy consciente de la estructura causal necesaria para que yo haga Y.

Sé que a veces hago Y.

Luego, soy consciente de que no se requiere ninguna estructura causal para que yo haga Y.

Lo que él defiende es que el argumento es inválido. No ser consciente de algo no es lo mismo que ser consciente de su ausencia. Cuando hablo, no soy consciente de la increíble estructura causal que lo hace posible: la musculatura, el control de la respiración, el movimiento de la lengua y el paladar, la configuración de mi mandíbula. Y sin embargo todo ello es necesario, tal como pronto descubriría si fallase alguna de estas cosas.

En este punto alguien podría pensar lo siguiente:

Si restringimos nuestros pensamientos al mundo físico, puede que nos parezca que no nos queda más opción que el determinismo o una indeterminación arbitraria y perdamos de vista la libertad real. Pero supongamos que existe otro nivel. En algún lugar, detrás de las evoluciones del cerebro y del cuerpo, o quizá por encima de ellas, está el Verdadero Yo, el cual recibe información y ocasionalmente dirige algunas operaciones. Habrá ocasiones en las que el cerebro y el cuerpo, dejados a su antojo, actuarían de una forma. Pero la dirección del Verdadero Yo hará que actúen de otra forma. Puedo tomar el control e interferir en el modo en que las cosas se habrían desarrollado si no lo hubiera hecho. En

esto consiste mi libertad.

En tal caso estamos describiendo una cierta relación entre mí mismo, por un lado, y mi cerebro y mi cuerpo, por el otro, como si fuera una interacción entre dos instancias distintas. El cerebro y el cuerpo transmiten mensajes al Verdadero Yo, y éste responde con instrucciones. El Verdadero Yo se halla en la cabina de mando, y la persona en conjunto se comporta libremente cuando éste tiene la situación bajo su control. Si no, el cerebro y el cuerpo siguen adelante con sus («inconscientes») evoluciones físicas.

De nuevo encontramos el dualismo mente-cuerpo. El Verdadero Yo dirige los acontecimientos. Los mensajes llegan hasta él, quizás a través de la glándula pineal. Un soplo del alma pone en marcha las neuronas y las sinapsis, e inicia nuevas cadenas causales. Hay un fantasma en la máquina, y la máquina se comporta libremente cuando el fantasma tiene el control. Ya hemos tenido ocasión de conocer algunos de los misterios de la interacción entre la mente y el cuerpo. Pero en este caso podemos proponer una objeción distinta. El dualismo trata de comprender la libertad humana mediante la introducción de un ingrediente extra, el alma. Pero ¿de dónde sale la libertad del alma?

Examinemos otra vez el dilema del determinismo. ¿Cómo escapa a este mismo problema un fantasma dentro de la máquina o un alma? ¿Acaso hay leyes que gobiernen el comportamiento del elemento espiritual, de modo que si un espíritu se halla en un determinado estado en un momento dado, existe una ley que determina cuál será su siguiente estado? Si no es así, ¿acaso el elemento espiritual está sometido al azar? ¿En qué me ayuda eso a ser libre y responsable? Recordemos también que no hay ninguna garantía divina de que lo mental, por el mero hecho de serlo, tenga que estar bajo mi control: no puedo desear sin más que desaparezcan los dolores, los deseos, las obsesiones, los pensamientos indeseables y las confusiones.

La concepción dualista de la libertad comete un error filosófico fundamental. Advierte un problema y trata de resolverlo llamando a escena a un nuevo tipo de «cosa». Pero se olvida de preguntar cómo escapa esta nueva «cosa» al problema que se había planteado en relación con las cosas normales. Encontraremos esta misma clase de error en el capítulo 5, al hablar de la filosofía de la religión. En realidad, si pensamos acerca de ello veremos que sin darnos cuenta concebimos la libertad de cualquier alma inmaterial, o de cualquier fantasma en la máquina, desde el modelo de la libertad humana. Es decir, que lejos de ayudarnos a comprender la libertad humana, la propia idea a la que hemos acudido depende de ella. En realidad, el fantasma no es más que una especie de pequeño ser humano etéreo, un «homúnculo» que recibe información, delibera, desea cosas, es influido o inspirado o guiado por ciertas informaciones, y eso es todo lo que podemos imaginar que hace. Si no somos capaces de comprender la libertad en los seres humanos, tampoco podemos comprenderla en el caso de un tal homúnculo.

Y por supuesto, queda en pie el problema de la relación entre la mente y el cuerpo, que tan intratable se ha mostrado en el marco del dualismo cartesiano. El sistema físico es un sistema cerrado. Se requiere una causa física para producir un efecto físico.

En un intento de reconciliar la idea de la libertad con un universo determinista compuesto por pequeños átomos en movimiento, sólidos e indivisibles, el filósofo *griego* Epicuro (341-270 a.C.) ya había sugerido que el espíritu de una persona podría intervenir para «desviar» la dirección de los átomos. Lucrecio, que en realidad está interpretando a Epicuro en el pasaje que aparece al comienzo del capítulo, habla más adelante de cierta desviación mínima que experimentan los átomos y de cómo esta «exigua desviación de los átomos no está determinada en el tiempo ni en el espacio». Lamentablemente, las leyes físicas del movimiento son poco propicias a esta «desviación». Las leyes descubiertas hasta el momento nos

dicen que la inercia lineal, una función combinada del movimiento y la dirección, se conserva. Algo así atentaría tanto contra las leyes del movimiento como si el Verdadero Yo pudiera cambiar la dirección de la Luna con sólo pensarlo, o como si pudiera acelerar o frenar su movimiento.

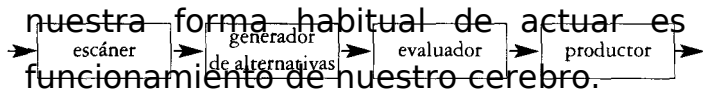
Al margen de esto, sin embargo, resulta interesante señalar que tanto los atomistas griegos como los romanos, incluidos Epicuro y Lucrecio, se encontraban en mejor posición que Descartes al menos en un aspecto. A diferencia de él, pensaban que el propio espíritu debía ser concebido en términos mecánicos. Sostenían que la mente o espíritu se compone de unas partículas mecánicas particularmente sutiles, muy pequeñas y extremadamente móviles, de modo que en principio nada impide que influyan en las direcciones y las velocidades de las partículas más grandes que componen un cuerpo. Lucrecio explica que este sutil elemento está formado por «partículas minúsculas, que extienden su trama por venas, vísceras y nervios». El alma tiene que estar hecha de un elemento tenue, ya que «sueños de humo y de niebla pueden moverla». Tales sueños estarán hechos presumiblemente de partículas más pequeñas aún que las del humo y la niebla. Por desgracia, Lucrecio tampoco resuelve la cuestión de cómo el movimiento de estas minúsculas partículas puede romper los lazos del destino y escapar a la cadena infinita de las causas y los efectos. A los antiguos atomistas les gustaba comparar la acción del alma sobre el cuerpo con la acción del viento sobre un velero, aunque por supuesto el viento forma parte de la cadena infinita de las causas y los efectos. No es algo que escape a ella, como tampoco lo es el alma en este modelo.

Dominarse a uno mismo

¿Hay alguna respuesta mejor al argumento del incompatibilismo?

El argumento del determinismo radical no dice nada acerca de los

diferentes tipos de relaciones causales que pueden estar en juego cada vez que un agente lleva a cabo una determinada acción. Sin embargo, a veces la cadena causal es enteramente independiente de lo que nosotros pensemos. La cadena causal que lleva de mi inmersión irreversible bajo el agua a mi muerte por ahogamiento es una de ellas. El resultado es tan inevitable en el caso de Einstein como en el de un burro. Pero a veces la cadena causal sigue la vía de los procesos neurales de alto nivel. Esto es tanto como decir que ~~nuestra forma habitual de actuar es el resultado del correcto funcionamiento de nuestro cerebro.~~



Probemos, pues, con un modelo más primitivo. Pensemos en el cerebro en términos de *software*, como si estuviera dividido en varios «módulos». Uno (un «escáner») le permite recibir información acerca de una determinada situación. Otro (un «generador de alternativas») define las opciones posibles de comportamiento a la luz de la información proporcionada por el escáner. Un tercer módulo (un «evaluador») establece el orden de prioridad de las diversas opciones en función de los objetivos para los que está programado. Puede que funcione a base de asociar distintos indicadores emocionales, como el miedo o la alegría, a cada una de las opciones. Finalmente, un cuarto módulo (un «productor») toma la primera opción seleccionada por los procesos anteriores y envía las señales neuronales pertinentes para poner en movimiento los músculos y los miembros del cuerpo. Este sería el esquema básico:

Recordemos que todo esto no es más que una descripción de las partes del cerebro en términos de *software*. Supongamos ahora que el funcionamiento de dichos módulos ha dado como resultado una determinada decisión. Supongamos que se trata de una decisión tuya, y que estos módulos han trabajado para producirla tal como lo hacen normalmente. Si les damos el nombre de «módulos de



decisión», y aceptamos su implicación en la elaboración de la respuesta, entonces se puede decir perfectamente que la respuesta la escogiste tú. No fue algo que se te impusiera, como por ejemplo el ahogarse es algo que se impone al bañista atrapado.

Supongamos que la decisión consistía en hacer algo realmente malo. Entrás en mi habitación y tiras a mi pacífico perro por la ventana. Yo me siento ultrajado, y me dispongo a tomar represalias. Supongamos que tratás de defenderte invocando el argumento del incompatibilismo.

Ten en cuenta que esta acción ha sido el resultado de la «configuración» de mi sistema de módulos de decisión. Quizás algunos sucesos de mi infancia, sobre los que tengo escaso control, lo «configuraron» de tal modo que conseguir un entorno sin perros tiene la máxima prioridad para mí. Mi generador de alternativas me dijo que había una opción, tras ser informado por el escáner de que estaba en presencia de un perro y de que había una ventana cerca. Mi evaluador seleccionó inmediatamente esa opción, y mi productor inició sin más dilación la acción de tirar al perro por la ventana. ¿Por qué me acusas de ello?

Es probable que yo no esté demasiado impresionado. Mi respuesta sería muy parecida a ésta:

No tengo el menor interés en saber cómo llegaste a estar «configurado» de este modo. Lo que me molesta es que ésa sea tu configuración. No me importa cómo llegó a serlo, o qué fuerzas deterministas hicieron que tus sistemas se configuraran de este modo. Lo único que me importa es que a fin de cuentas el resultado es odioso, y por eso voy a darte una paliza. Quizá fue pura mala suerte que acabaras siendo así. Y ahora la mala suerte será doble, porque vas a recibir una paliza por ello.

¡Siempre me queda el consuelo de saber que, según tu propio argumento, no me puedes culpar por el hecho de pegarte! Es sólo la

forma en que estoy configurado: reacciono mal cuando la gente le hace estas cosas a mi pacífico perro.

El hecho de darte una paliza puede tener una ventaja —en realidad, varias ventajas—. Puede que sirva para reajustar tu evaluador. La próxima vez, puede que este módulo ponga la opción de tirar el perro por la ventana por debajo de la opción de soportar su presencia en el orden de prioridades. También podemos imaginar mecanismos más complicados que tengan ese mismo efecto —tal vez ponga una señal en la opción «tirar al perro» que diga: «Riesgo de recibir una paliza»—. O quizá mi furia haga que te replantees de modo más general tus estrategias de comportamiento. E incluso si el hecho de pegarte no tiene ningún efecto sobre ti, al menos envía una señal a otros posibles enemigos de los perros. Eso también representa un alivio.

Hacer esto no es lo mismo que culpar a alguien por haberse ahogado, por ejemplo, y en cambio no culparle por haber quedado atrapado bajo el agua. En este último caso, la cadena causal depende de la fisiología básica de los animales, algo que no puede cambiarse a través de la educación o de la actitud de los demás. No puede ser «reconfigurado» por medio de premios o castigos. La cadena causal no depende en este caso de módulos elásticos o flexibles, es decir, que puedan reconfigurarse por aquellos medios. En cambio, sí podemos disuadir o reeducar a aquellos que tiran perros.

Los profesores de escuela dicen a veces cosas como ésta: «No me importa tener un alumno estúpido, pero sí me molesta tener uno perezoso». Llevados por el argumento del determinismo radical, podríamos pensar que esto no es más que un prejuicio: algunas personas nacen estúpidas y sentimos compasión por ellas; ¿por qué no deberíamos sentir compasión también por aquellos que han nacido perezosos? No es más que una cuestión de suerte. Pero la actitud del profesor de escuela tendría un sentido si la pereza, a diferencia de la estupidez, respondiera a ciertos incentivos. Si el respeto por la

opinión de nuestro profesor nos hace trabajar más duro, aunque no nos haga más inteligentes, entonces la asimetría está justificada. El profesor tiene la función de reconfigurar nuestro módulo evaluador. La experiencia humana demuestra empíricamente hasta qué punto pueden llegar a reconfigurarse los módulos de decisión en las relaciones con los demás, incluidas las relaciones desagradables como el enfado o el desprecio.

Aquí tenemos un primer esbozo —pero sólo un esbozo— del programa del compatibilismo: el intento de mostrar que, si se plantea el problema correctamente, no hay ninguna inconsistencia entre el determinismo y nuestra práctica habitual de atribuir a las personas la responsabilidad de sus acciones. A veces recibe el nombre de determinismo «moderado», en oposición al determinismo radical. No resulta una etiqueta demasiado acertada, y ello por dos razones. Primero, no se trata realmente de una clase diferente de determinismo. Lo asume exactamente en el mismo sentido que todo el mundo: no hay ningún poder espiritual que venga a interferir en el orden causal natural de los acontecimientos. Segundo, en términos morales o políticos, el determinista «moderado» puede llegar a ser bastante radical, en el sentido de duro. Sí alguien se le acerca con la melodramática excusa de que su biología o su entorno le ha hecho tal como es, se vuelve sordo y descarga igualmente su furia. La fácil ecuación entre el crimen y la enfermedad no tiene ningún sentido para él: la gente es muy capaz de hacer lo que debe hacer, y en caso contrario empleará el castigo o cualquier otro recurso apropiado para obligarles a hacerlo.

Por supuesto, el compatibilista también admite cierta clase de excusas. Si una persona estuviera tan condicionada en una determinada situación que le resultara imposible encontrar una buena salida, por mucho que sus «módulos» estuvieran en perfecto funcionamiento, no sería culpable de lo que pudiera suceder. Tal es el caso del bañista que se ahoga: no puede hacer nada por salvarse, por

mucha personalidad que tenga. De modo parecido, si alguien se ha visto obligado a tomar, por ejemplo, una medicación que provoca desorientación o un estado depresivo, tal vez se le puedan perdonar sus acciones mientras se encuentre en este estado.

Llegados a este punto es posible que pensemos: bueno, la reacción frente al malvado enemigo de los perros fue bastante natural. Tal vez resulte incluso justificable en términos de sus consecuencias. Tal vez la idea de culpa y las reacciones asociadas a ella simplemente cumplan una función necesaria. Pero a pesar de ello, ¿acaso no descubrimos un indicio de injusticia? No hemos hecho nada para demostrar que el enemigo de los perros podría haber actuado de otro modo. El resultado de nuestras acciones está siempre predeterminado por la configuración de nuestros módulos. Por lo tanto, los compatibilistas parecen responsabilizar a las personas de acciones que éstas no podían evitar. Sin embargo, ellos podrían replicar distinguiendo varios sentidos de la expresión «podría haber actuado de otro modo». Si la cadena causal que lleva a una determinada acción se sitúa en los módulos de decisión del agente, podemos decir que en cierto sentido «podría haber actuado de otro modo» y, por lo tanto, sería libre. Para comprender adecuadamente lo que significa «podría haber actuado de otro modo», podemos fijarnos en lo que yo llamaría el *primer principio del compatibilismo*:

Podemos decir que un sujeto ha actuado libremente si hubiera podido actuar de otro modo en sentido propio. Esto significa que el sujeto *hubiera* actuado de otro modo *si* hubiera tomado una decisión diferente.

Y de acuerdo con el compatibilista esto justifica de modo suficiente el hecho de que consideremos responsables a las personas y tal vez que descarguemos sobre ellas nuestra furia e indignación.

La réplica espiritualista frente al determinismo propuso una intervención exterior al reino de la naturaleza: una libertad «no causal». El espíritu estaría fuera del orden causal de la naturaleza y,

sin embargo, sería misteriosamente capaz de alterarlo. Podemos llamar a esta concepción *control intervencionista*. En algunos textos recibe el nombre de concepción libertaria de la libertad, aunque eso puede llevar a confusiones, ya que no tiene ninguna relación con el liberalismo político o económico, que consiste en la ideología del libre mercado y el mínimo gobierno. Voy a seguir refiriéndome a ella como control intervencionista. El compatibilismo, en cambio, considera que la persona está perfectamente ubicada dentro del orden causal de la naturaleza. Su libertad consiste en que sus acciones dependen de sus propios procesos cognitivos. Entonces, ¿cómo responde el compatibilista al argumento original sobre el control? Acaso sugeriría que no necesitamos más argumento que éste:

El pasado controla el presente y el futuro.

Un termostato no puede controlar el pasado.

Un termostato no puede controlar la forma en que el pasado controla el presente y el futuro.

Luego, un termostato no puede controlar el futuro.

Algo debe fallar en este argumento, ya que un termostato sí puede controlar el futuro, en lo que se refiere a la temperatura. Esto es precisamente lo que hacen los termostatos. Un termostato controla la temperatura gracias a que forma parte de la cadena causal que va desde el pasado hacia el presente y el futuro. Y de acuerdo con el compatibilismo, así es como controlamos el futuro. Estamos implicados en el orden causal. Somos parte de la cadena que va del pasado al futuro. Y éste es el origen de nuestra responsabilidad. Podemos llamar a esta concepción «control dentro del control», o «control desde dentro de la naturaleza». Cuando ejercemos este control interno, defiende el compatibilista, somos responsables de ciertos sucesos. Y si ejercemos mal ese control, está perfectamente justificado que nos consideren responsables del resultado, y que seamos condenados por ello si procede.

Pero ¿acaso es ésta la libertad que buscábamos? No atribuimos ninguna libertad a un termostato. Esta es la razón por la que algunos han visto en el compatibilismo una forma de ignorar el problema de la libertad, antes que una forma de solucionarlo. Tal era la opinión del gran Immanuel Kant (1724-1804), quien lo rechazó por considerar que no nos daba más que la «libertad del reloj» y lo tachó nada menos que de «perverso subterfugio».

Títeres y marcianos

Todavía hay otra razón para compartir la opinión de Kant. Los módulos y las complejidades del procesamiento de la información han acabado por complicar el escenario causal. Pero ¿acaso lo alteran en lo fundamental? Imaginemos un alegato en favor de la higuera que señalara que en aquel momento era invierno y no verano. Eso sería una perfecta defensa para la higuera. Pues bien, ¿acaso no puedo decir también que es invierno cada vez que actúo mal? Mis módulos habrían sido mal configurados, presumiblemente a causa de sucesos pertenecientes a cadenas causales que se extienden hasta mucho antes de mi nacimiento. Tal vez el hecho de que alguien se enfade conmigo pueda alterar mi sistema de toma de decisiones en el futuro, pero no demuestra que yo hubiera podido actuar de otro modo en el pasado.

A medida que vamos descubriendo las regularidades causales que se esconden detrás de las decisiones y otros estados mentales, es probable que pasemos a planteamientos menos moralistas. Podríamos reprocharle a alguien su tendencia a estar deprimido todo el tiempo, hasta que descubriéramos el historial químico que lo explica. Podríamos estar enfadados con alguien por ser incapaz de mover un dedo, hasta que supiéramos que padece mononucleosis. Pero, según el determinista, siempre hay cosas como éstas por descubrir. Sin necesidad de acudir a otros ejemplos neurofisiológicos,

podemos pensar simplemente en el caso de un «lavado de cerebro» o «condicionamiento». A menudo los padres reprenden a sus hijas adolescentes por gastar tiempo, energía y dinero en cosméticos perfectamente prescindibles, pero serían más comprensivos si tuvieran en cuenta las presiones sociales y comerciales que bloquean su buen juicio y las llevan a esta situación.

Las cosas se ponen todavía más difíciles para el compatibilismo si nos permitimos un poco de ciencia ficción. Imaginemos la invasión de unos marcianos increíblemente pequeños, maliciosos y bien organizados: lo bastante pequeños como para invadir nuestros cerebros y pasearse por dentro. En tal caso, los marcianos podrían configurar nuestros módulos prácticamente a su antojo. Estaríamos en sus manos, como si fuéramos títeres. (Si este tipo de ejemplo suena demasiado rebuscado, téngase en cuenta que de hecho existe un parásito cuya estrategia vital consiste en colonizar el cerebro de las hormigas. Bajo su influencia, la hormiga tiende a subirse a las hojas de hierba. Esto hace más probable que sean ingeridas por alguna oveja que pase por allí, la cual resulta a su vez infectada por el parásito [el individuo concreto que se halla en el cerebro de la hormiga perece, pero otros se suben al carro]. Que nosotros sepamos, la hormiga se siente perfectamente libre mientras se encarama a su hoja de hierba.) Por supuesto, podría ser que los mini marcianos nos pusieran a hacer lo mismo que haríamos nosotros en cualquier caso. Pero también podría ser que pulsaran las teclas químicas adecuadas para que hiciéramos cosas terribles. Supongamos entonces que la ciencia, por fortuna, inventa un detector que nos permite saber si los marcianos nos han invadido. ¿Acaso no seríamos compasivos con cualquiera que hubiera sufrido tal desventura? ¿Acaso no reconoceríamos inmediatamente que no era responsable de sus malas acciones?

Sin embargo, diría el incompatibilista, ¿acaso importa si la culpa fue de los marcianos o de alguna causa de tipo más natural?

Este tipo de respuesta pone en entredicho la interpretación compatibilista de la expresión «no podría haber actuado de otro modo». Está muy bien, viene a decir, afirmar que alguien podría haber actuado de otro modo si hubiera tomado una decisión diferente. Pero supongamos que estuviera configurado de modo que no pudiera escoger otra cosa. Supongamos que en el momento de actuar sus módulos de decisión estuvieran bloqueados en cierta posición por los marcianos, la química o lo que sea. ¿Y entonces qué? El compatibilista que hemos visto hasta ahora se quitaría el problema de encima: no le interesa saber cómo hemos llegado a ser como somos, sólo si el resultado es bueno o malo. Su oponente cree que sí es importante, y por lo menos algunas de nuestras reacciones al conocer más cosas acerca de las cadenas causales demuestran que estamos de acuerdo con él.

Obsesiones y Twinkies

Creo que la mejor línea de argumentación que puede emplear el defensor del compatibilismo frente a este contraataque es cuestionar el significado de la palabra «configuración», cuando decimos que los módulos están configurados para producir cierto resultado. Se trata de una estrategia parecida a la que había empleado antes al distinguir el hecho de tomar una decisión del hecho de ahogarse. En aquella ocasión introdujo un cierto grado de flexibilidad en el proceso causal al señalar que los módulos pueden ser ajustados o configurados de un modo distinto. Cuando su oponente afirma otra vez que el sujeto no es más que una víctima de unos módulos mal «configurados», la réplica debería consistir en introducir otro nivel de flexibilidad. Es cierto, podemos decir: tanto en el caso de la adolescente como en el de los marcianos puede ser que la configuración de los módulos haya sido realmente predeterminada, sea a causa de procesos químicos o de otro tipo. Pero se trata de

casos especiales, precisamente porque cuando el sujeto se encuentra en esta clase de situaciones se vuelve inflexible: inmune a la argumentación o a las novedades o cambios que se puedan producir en el escenario de la toma de decisiones. Pero normalmente los agentes no están tan predeterminados en su comportamiento. Su libertad consiste en el hecho de que son receptivos ante nuevas informaciones y ante las diferencias que puedan darse en la situación. No se ven empujados de forma irresistible a tirar perros por la ventana o a quedarse todo el día en el mostrador de los cosméticos.

Podríamos desarrollar esta idea del siguiente modo, que llamaré el *principio revisado del compatibilismo*:

Podemos decir que un sujeto ha actuado libremente si hubiera podido actuar de otro modo en sentido propio. Esto significa que el sujeto hubiera actuado de otro modo si hubiera tomado una decisión diferente y que, en caso de hallarse bajo la influencia de otros pensamientos o consideraciones, habría tomado una decisión diferente.

Por supuesto, en alguna ocasión, el hecho de que no se susciten los pensamientos adecuados se deberá a la mala suerte. Pues bien, dirá una vez más el compatibilista: «Eso será toda la mala suerte que quieras. Pero quizás mi furia y el hecho de que vaya a darte una paliza evitará que se vuelva a repetir».

Algunos filósofos (Baruch Spinoza [1632-1677] es el ejemplo más conocido) tienden a asociar la libertad con un mayor grado de conocimiento y comprensión de las cosas.³ Según ellos, somos libres en la misma medida en que somos sabios. La idea resulta atractiva por muchos motivos: vincula la libertad de la voluntad a cosas como las libertades políticas (libertad de información y libertad de expresión). Sólo somos libres en la medida en que tenemos oportunidades abiertas ante nosotros, y la falta de información nos niega estas oportunidades. Podríamos añadir esta reflexión al principio revisado del compatibilismo, especificando que los «otros

pensamientos o consideraciones» son, en primer lugar, representaciones adecuadas de la situación en la que se encuentra el agente y de las opciones que tiene y, en segundo lugar, que deben ser asequibles para él. Dicho de otro modo, no sirve de mucho asegurar que bajo la influencia de otros pensamientos o consideraciones el sujeto habría tomado una decisión diferente, si esos otros pensamientos y consideraciones simplemente no entraban en el juego. Así, supongamos que me propongo envenenarte y que astutamente pongo arsénico en tu café. Lo bebes. No sirve de mucho decir que eras libre de no hacerlo, pues aunque es cierto que podrías haber rechazado el café si así lo hubieras decidido, y también que el pensamiento o consideración de que el café podría estar adulterado con arsénico habría hecho que tomaras una decisión diferente, sin embargo, como no había razón alguna para que tal idea se te pasara por la cabeza, fuiste más una víctima que un agente libre. Podríamos incorporar este punto a una versión revisada del principio revisado del compatibilismo:

Podemos decir que el sujeto ha actuado libremente si hubiera podido actuar de otro modo en sentido propio. Esto significa que hubiera actuado de otro modo si hubiera tomado una decisión diferente y que, en caso de hallarse bajo la influencia de otros pensamientos o consideraciones verdaderos y asequibles, habría tomado una decisión diferente. Son pensamientos y consideraciones verdaderos y asequibles para el agente aquellos que representan adecuadamente la situación en la que se encuentra y que sería razonable esperar que tomara en consideración.

¿Qué pasa cuando a una persona simplemente no se le ocurren aquellos pensamientos o consideraciones? ¿Acaso es más una víctima que un agente responsable? Ello introduce un nuevo giro en la situación.

Hasta ahora hemos hablado como si la «libre elección», sea

procedente de alguna misteriosa intervención exterior o de algún «interior» sustitutivo de tipo compatibilista, fuera un requisito necesario para la responsabilidad. Pero ¿es eso cierto? Tal como he afirmado antes, el hecho de que una persona no se plantee algún pensamiento crucial a la hora de tomar una decisión podría ser una cuestión de simple mala suerte. Pero a veces no lo tratamos como «simple» mala suerte. Decimos que se le debería haber ocurrido. El agente se expone a ser censurado si no se le ocurre. Alguien que se divierta provocando incendios en edificios de viviendas no puede alegar seriamente que «nunca se le ocurrió» que alguien podía resultar herido, a menos que sea un niño o un deficiente mental. Incluso si fuera verdad que no se le había ocurrido, y no hubiera, por lo tanto, una decisión libre de poner a terceras personas en peligro, seguiría siendo responsable. Tanto la temeridad como la negligencia son faltas y somos tan responsables por ellas como lo podemos ser por decisiones más controladas. Eso ha resultado difícil de aceptar para algunos filósofos. Aristóteles sostuvo el dudoso principio de que las personas negligentes en realidad habían decidido convertirse en negligentes, tal vez en algún estadio temprano de la infancia, y que ésta era la única razón por la que podíamos considerarlas responsables.⁴

En este punto se abre ante nosotros un amplio abanico de reflexiones interesantes. Algunos tipos de mala suerte son realmente contingentes y, por lo tanto, no afectan a nuestra relación con el agente. Otros, en cambio, hacen que nos planteemos algunas reflexiones acerca de él. Imaginemos un golfista. Supongamos que el primer día da un buen golpe, pero, increíblemente, una gaviota se interpone en la dirección de la bola e impide que llegue al lugar deseado. El segundo día da un golpe igualmente bueno, pero una ligera brisa desvía la bola y vuelve a impedirlo. Podríamos decir que en ambos casos se trata de mala suerte. El primero es pura mala suerte, pero el segundo no es tan sencillo. Se trata de mala suerte, sí,

pero es la clase de mala suerte que se espera que un golfista realmente bueno sepa prever y evitar con su juego. Es algo que debería estar al alcance del golfista. En cambio, el caso de la gaviota representa un puro acto de Dios. A partir de cierta cantidad de mala suerte de la segunda clase, nuestra opinión del golfista empeora, y eso mismo es lo que sucede en todos los demás casos. Eso explica la respuesta de un pianista a un admirador suyo que se maravillaba de la suerte que tenía por poseer tanto talento: «Sí y, cuanto más practico, más suerte tengo».

El trabajo de ingeniería conceptual que estamos realizando pretende desenmascarar o hacer explícitos ciertos aspectos de nuestra forma de pensar. Queremos resaltar y sintetizar en la medida de lo posible cuestiones como las siguientes: reconocemos una diferencia entre cambiar el pasado (algo que no podemos hacer) y actuar de modo distinto a como lo hacemos (algo que a veces podemos hacer); nuestras prácticas en relación con la culpa son selectivas; reconocemos una diferencia entre estar enfermo y ser salvado; consideramos admisibles ciertas excusas, mientras que otras no. El análisis filosófico pretende darnos un dominio intelectual sobre todas estas cuestiones. Pretende demostrar que son el resultado de la aplicación de un conjunto razonable y defendible de conceptos y principios, y no sólo una mezcla irracional de hábitos inconexos. La dificultad que entraña este proyecto hace que la filosofía sea una disciplina difícil. La propuesta compatibilista constituye un proyecto de ingeniería que pretende o bien describir la estructura de nuestros conceptos actuales, o bien diseñar nuevos y mejores conceptos. Tiene que ofrecer una explicación de nuestra forma habitual de pensar, o de aquella forma de pensar que nos permite controlar mejor nuestros problemas. Yo, personalmente, creo que la versión revisada del principio revisado del compatibilismo lo consigue bastante bien. Pero otros se toman más en serio la objeción kantiana. Piensan que nuestras «relaciones interpersonales», incluida

nuestra forma de atribuir responsabilidad a los demás o a nosotros mismos, se basan en que conservamos cierta confianza en un modelo de libertad de tipo intervencionista. De modo que si esta última posición está desahuciada desde el punto de vista metafísico, deberíamos revisar nuestras prácticas. El problema filosófico sería entonces que el control intervencionista resulta insostenible, y el control interno inadecuado.

A veces el análisis nos permite resolver los casos difíciles. Pero otras veces deja algunos puntos oscuros, y ello no tiene por qué ser malo. Volvamos a la chica adolescente que dedicaba una cantidad increíble de tiempo y dinero a los cosméticos. ¿Puede actuar de otro modo? Si aplicamos la versión revisada del principio revisado del compatibilismo, veremos que la cuestión depende ahora de qué otros pensamientos o consideraciones le resultan «asequibles». Por un lado, nos sentiríamos inclinados a decir que podría comenzar por darse cuenta de que su popularidad o su atractivo no ganan mucho con los cosméticos (ganarían mucho más si aprendiera a pensar de una forma mínimamente presentable, quizá gracias a la lectura de un libro como éste). Tal vez pensemos que esta idea resulta de lo más acertada y que está potencialmente al alcance de todo el mundo. Pero en cierto sentido es posible que no sea así. Puede que las personas sometidas al tipo de influencias a las que está sometida ella sean simplemente incapaces de considerarla seriamente. Nuestra cultura tiene una extraordinaria capacidad para hacer que los adolescentes sean sordos a esta verdad, de modo que no sería razonable esperar de ella que la considerara seriamente. Personalmente, me inclinaría por este último diagnóstico, y la vería más como una víctima que como un sujeto libre. Pero lo importante es que aunque la versión revisada del principio revisado no resuelve esta cuestión, ciertamente llama la atención sobre ella. Y eso mismo ya es un paso adelante a la hora de dominar la cuestión de la responsabilidad y la libertad. Pero es justo añadir que queda todavía

mucho camino por delante. Un incompatibilista, por ejemplo, podría insistir en que los pensamientos son asequibles sólo en la medida en que son objeto de una selección libre (e intervencionista), y ello volvería a situarnos en un empate.

La cultura contemporánea no ayuda a tener una idea demasiado clara de lo que es la responsabilidad. Pensemos en la célebre «defensa Twinkie». Un día de 1978, un ex empleado del ayuntamiento de la ciudad de San Francisco, Dan White, entró armado en el edificio del ayuntamiento a través de una ventana del sótano, evitando así el detector de metales. Subió a los pisos de arriba, disparó y mató al alcalde George Moscone y al supervisor Harvey Milk. En el juicio, el psiquiatra de la defensa, Martin Blinder, testificó que White había sufrido una depresión, lo cual le llevó a comer demasiado, en particular el tipo de comida basura altamente azucarada conocida como Twinkies. Según la versión de Blinder, ello no hizo más que agravar su depresión, ya que White era un ex atleta y sabía que los Twinkies no eran buenos para su salud. Blinder sostuvo que White debía encontrarse en un estado emocional que hacía imposible que actuara con premeditación o verdadera intencionalidad, dos agravantes necesarios para apreciar asesinato en primer grado. El argumento impresionó al jurado, que absolvió a White del cargo de asesinato y le condenó, en cambio, por el crimen de «homicidio voluntario», de menor gravedad.

Después de este caso, el Estado de California revisó sus leyes para limitar el margen que ofrecían a esta clase de defensas y, según parece, tenía buenas razones para hacerlo. White actuó obviamente con intencionalidad y premeditación —por eso adquirió un arma y entró por el sótano—. Y es fácil comprobar que la versión revisada del principio revisado no es nada propicia a la defensa Twinkie. Un acusado tendría que emplearse a fondo para demostrar que una cantidad excesiva de azúcar arranca literalmente nuestro comportamiento de la esfera de nuestros módulos de decisión y de

nuestros pensamientos. No parece razonable decir que con una cantidad suficiente de Twinkies en el cuerpo nos volvemos literalmente incapaces de tener ciertos pensamientos, de modo que no se podría esperar razonablemente de nosotros que cayéramos en la cuenta de que asesinar a personas es una mala idea, por ejemplo. Ni siquiera una montaña de azúcar podría conseguir algo parecido. (Hay que decir que los jurados contemporáneos tampoco tienen una idea demasiado clara de lo que es la causalidad. Recientemente, un hombre ganó un litigio por daños morales en Michigan porque, según alegó, una colisión trasera con su coche le había convertido en homosexual.)

Antes de abandonar el compatibilismo, vale la pena señalar un problema al que se enfrentan todos los principios que hemos visto hasta ahora. El compatibilismo trata de construir el concepto adecuado de control a partir de la idea de que si las circunstancias hubieran sido diferentes, el agente habría actuado de otro modo. Pero algunos casos difíciles sugieren que estas ideas no encajan de una forma tan perfecta: son casos en los que existe lo que se llama una *sobredeterminación causal*. En tales casos hay algo que ejerce un control efectivo sobre el resultado de un determinado proceso, a pesar de que el resultado hubiera sido el mismo si no lo hiciera, gracias a un cierto mecanismo de seguridad. Así, podría ser que un termostato estuviera controlando la temperatura, a pesar de que ésta hubiera sido la misma si el termostato no hubiera funcionado, gracias a un mecanismo de seguridad. Si el termostato no hubiera funcionado, algún otro mecanismo se hubiera puesto en marcha para mantener la temperatura al nivel adecuado. De modo parecido, alguien podría cometer una mala acción de forma intencionada, responsable y controlada, a pesar de que si hubiera querido actuar de otro modo, algún mecanismo desconocido se habría puesto en marcha para asegurar que cometía igualmente la mala acción. Imaginemos que los marcianos estuvieran escondidos en su sitio, sin

interferir por el momento en el desarrollo de los acontecimientos, pero dispuestos a hacerlo en caso de que pareciera que el resultado iba a ser distinto del que ellos querían. Tales casos resultan extrañamente difíciles de resolver. Sin embargo, el compatibilista puede considerar que no añaden ninguna dificultad a la definición del concepto adecuado de control en el caso de los seres humanos que no se pueda plantear también en el caso de los termostatos. Y como el problema tiene necesariamente una solución para el caso del control mecánico, debe tenerla también para el caso de las personas.

Objetivar a las personas

¿Hay algún otro problema que deba preocuparnos? Alguien podría pensar del siguiente modo:

El planteamiento compatibilista describe el funcionamiento de los seres orgánicos con cerebro en términos de módulos de decisión. Pero esto no es más que describir las cosas en términos de lo que pasa. No es lo mismo que describir las cosas en términos de la actividad de un sujeto, de lo que hago yo. Significa, por lo tanto, olvidar algo esencial en relación con mi propia humanidad, y esencial para el respeto que me merecen las demás personas como seres humanos: el hecho de que no somos víctimas pasivas, sino sujetos activos.

Así es como nos contemplamos a nosotros mismos y a las demás personas normales, y tal como normalmente desearíamos que nos vieran.

Así pues, el peligro es que estemos olvidando algo esencial para la vida humana. Es esencial que nos consideremos sujetos activos, y no víctimas pasivas. Y es esencial para nosotros el hecho de que los demás nos vean así. En un famoso artículo, el filósofo Peter Strawson (n. 1919) contrapone una actitud «objetiva» o impersonal hacia las demás personas con una actitud «personal» o humana.⁵ Desde la perspectiva objetiva, las demás personas no son más que obstáculos

en nuestro camino, que deberán ser «comprendidos, gestionados, tratados o controlados». No son objeto de actitudes personales. Tratamos a las demás personas como si fueran perturbados, antes que como sujetos inteligentes que pueden ser comprendidos.

Strawson propone un interesante cambio de perspectiva. En un principio, podría parecer que las actitudes morales asociadas con la culpa son severas y brutales, y podríamos ver como un progreso la posibilidad de que dieran paso a actitudes más liberales y comprensivas en relación con cosas como el crimen o el «comportamiento anormal». Tratar a las personas como si fueran pacientes antes que criminales parece ser un paso adelante en una dirección humana y decente. Strawson nos propone que pensemos en aquello que perdemos con el cambio: según él, buena parte de lo que distingue las relaciones humanas como tales. Supongamos, por ejemplo, que me he comportado de cierta forma y quiero dar explicaciones, pero me doy cuenta de que la actitud de los demás al escucharme sugiere que mi historia no es más que otro síntoma para ellos, no es más que otro indicio de que debo ser comprendido, gestionado, tratado o controlado. En tal caso, su actitud me habría deshumanizado. Quiero que comprendan mi decisión, no que la miren con condescendencia. Quiero que los demás «oigan mi voz», es decir, que valoren mi punto de vista, mi forma de ver las cosas, y no que se pregunten por la causa que puede hacer que un organismo humano se comporte así. Volveremos a encontrar este tipo de objetivación en el capítulo 8, al ocuparnos de la industria terapéutica, entre otras cuestiones.

La respuesta adecuada a la objeción que plantea Strawson sería la siguiente: el compatibilista no pretende negar la acción del sujeto, sino dar una determinada descripción de ella. Dicha descripción se plantea en términos de funciones modulares del cerebro, que consisten en introducir los datos, generar alternativas, ordenarlas según prioridades y, finalmente, dar una determinada respuesta, con

la que comienza la acción. Es cierto, tales procesos son cosas que «simplemente pasan» (de una forma pasiva, por así decirlo), pero según el compatibilista eso es lo que pasa, y todo lo que pasa, cuando una persona hace algo. Decir que alguien hace algo, y que lo hace por una determinada razón, es una descripción a nivel personal del resultado de todos aquellos procesos que tienen lugar a pequeña escala.

A algunos pensadores les gusta decir que hay dos perspectivas posibles acerca de este problema. Por un lado tenemos la actitud deliberativa, en primera persona, que adoptamos cuando somos nosotros los que tomamos una decisión. Y por otro lado tenemos una actitud «objetiva», o en tercera persona, que es la que adoptaría un científico al contemplarnos como un sistema neurofisiológico complejo y predeterminado. El problema es encontrar el modo de reconciliar ambas posturas.

Así planteado el problema, la solución correcta es seguramente la siguiente. Sólo tendríamos problemas para reconciliar ambas posturas si aquello que pone de manifiesto la actitud deliberativa fuera incompatible con aquello que pone de manifiesto la actitud objetiva. Pero la actitud deliberativa no manifiesta nada acerca de la causalidad. Pensar de otro modo es cometer el mismo error que cometió el agua en el ejemplo de Schopenhauer: confundir el hecho de no ser consciente del funcionamiento de la mente y el cuerpo con el hecho de ser consciente de la ausencia de tal funcionamiento. Lo primero se da de forma universal, mientras que lo segundo es imposible, puesto que sin tal funcionamiento no hay conciencia posible.

Así pues, como no hay nada en la actitud deliberativa que entre en conflicto con la concepción científica del mundo, tal vez no haya razón para encontrar difícil aquella reconciliación. No quedaría más que un problema moral: el de tomar conciencia de que no debemos tratarnos los unos a los otros desde una actitud objetivadora, sino

desde una plena comprensión humana, para la cual el conocimiento de las condiciones que influyen en las decisiones de los demás representa más un enriquecimiento que una limitación.

El destino, los oráculos y la muerte

Conocí a un hombre que había sido oficial durante la Primera Guerra Mundial. Me contó que uno de sus mayores problemas había sido conseguir que sus hombres llevaran el casco cuando se hallaban expuestos al fuego enemigo. El argumento de aquellos hombres iba sobre balas «que llevaban tu número escrito». Si una bala llevaba tu número escrito, no tenía sentido tomar precauciones, pues igualmente te iba a matar. Por otro lado, si ninguna bala llevaba tu número escrito, estabas a salvo un día más y no necesitabas llevar el incómodo y fastidioso casco.

Hay quien llama a este argumento el «sofisma perezoso». Si voy a tener un cáncer, pues voy a tenerlo, dice el fumador. No se puede escapar al destino. Y si el determinismo está en lo cierto, ¿acaso no está atado el futuro por la cadena interminable de los acontecimientos del pasado? El futuro nace de ellos: está escrito inevitablemente en el curso del tiempo. Y si el futuro ya está escrito, ¿acaso no deberíamos resignarnos a nuestro destino? ¿Acaso no pierde la acción su sentido? ¿Acaso no es mejor retirarse y quizá sentarse con un chal naranja diciendo «Om» todo el día?

Existen muchas historias que nos recuerdan la imposibilidad de escapar al destino. Veamos una versión de la famosa parábola islámica *Cuando la Muerte llegó a Bagdad*:

El discípulo de un sufí de Bagdad estaba un día sentado en un rincón de una posada, cuando oyó hablar a dos personajes. Por lo que decían, se dio cuenta de que uno de ellos era el Ángel de la Muerte.

«Tengo varias visitas que hacer en esta ciudad durante las próximas tres semanas», le decía el Ángel a su compañero.

Aterrorizado, el discípulo se escondió hasta que ambos hubieron partido. Para escapar a la Muerte, alquiló el caballo más veloz disponible y lo espoleó día y noche en dirección a la lejana ciudad de Samarcanda.

Mientras tanto la Muerte se encontró con el maestro sufí y hablaron sobre diversas personas. «¿Y dónde está tu discípulo tal y tal?», preguntó la Muerte.

«Debe de estar en algún lugar de esta ciudad, empleando su tiempo en contemplación, quizás en una posada», dijo el maestro.

«Qué extraño», dijo el Ángel, «pues se halla en mi lista. Sí, aquí está: tengo que recogerlo dentro de cuatro semanas, nada menos que en Samarcanda.»⁶

El discípulo intenta escapar a su destino, pero le alcanza igualmente. La historia de la huida imposible resuena por todo el mundo. En la tragedia *Edipo Rey*, de Sófocles, el rey Layo de Tebas escucha la profecía de que su hijo matará a su padre y se casará con su madre. Cuando tiene un hijo, Edipo, Layo trata de escapar al destino y lo abandona a su suerte, lisiado, en el monte. Pero Edipo es rescatado por un pastor y crece en Corinto, convencido de ser el hijo del rey de aquella ciudad. Oye rumores acerca de su destino y consulta el oráculo de Delfos, que lo confirma, de modo que huye en dirección contraria a Corinto, donde cree que se encuentra su padre. Y así, en un lugar del bosque donde se encuentran tres caminos, encuentra a Layo... El doble intento de frustrar el destino es precisamente lo que hace que se convierta en realidad.

Los soldados de mi amigo pensaban que tomar precauciones en su caso era tan inútil como la huida de Edipo de su desgracia. Pero hay una diferencia crucial. Se supone que Edipo conocía su destino, aunque intentaba evitarlo igualmente. En cambio, los soldados no sabían si morirían aquel día o no. Eso habría de permitirles encontrar la respuesta correcta, que consiste en decir que si una bala lleva o no tu nombre escrito puede ser que dependa de si decides llevar casco. Una bala que de otro modo hubiera llevado tu nombre escrito puede

seguir sin nada escrito con sólo tomar esta simple precaución. Y desde el momento en que no sabes si hay alguna bala que lleve tu nombre escrito, y te gustaría que ninguna lo llevara, sería mejor que tomaras la precaución.

No hacer nada —dejar de ponerse el casco, ponerse un chal naranja y decir «Om»— constituye una elección. Tener los módulos de elección configurados de acuerdo con el sofisma perezoso significa tener tendencia a esta clase de elección. Podemos formular el sofisma perezoso del siguiente modo:

El futuro será el que será. Lo que haya de suceder ya está escrito en el curso del tiempo.

Luego, no hagas nada.

Pero, ¿por qué habríamos de dejarnos impresionar por este argumento antes que por este otro?

El futuro será el que será. Lo que haya de suceder ya está escrito en el curso del tiempo.

Luego, ponte en marcha.

Sería preferible el primer argumento si supiéramos que las acciones humanas no tienen la menor influencia en el curso de los acontecimientos a lo largo del tiempo. Sería como si contempláramos desde detrás de una pared de espejo un juego en el que nunca podríamos participar, y cuyos jugadores fueran sordos y ciegos para nosotros. Pero las cosas no son así normalmente. Los acontecimientos siguen su curso, efectivamente, pero en secuencias harto predecibles. El hecho de que alguien se coma una tortilla viene siempre precedido por el hecho de que alguien rompa un huevo. El hecho de que alguien alcance la cima de una montaña viene siempre precedido por el hecho de que alguien se pone en camino. No hacer nada significa invariablemente que no se come ninguna tortilla o que no se alcanza ninguna cima. Lo que vaya a suceder depende de lo

que nosotros hagamos: esto es lo que significa la idea del control interno que ejerce una persona o un termostato. Nuestros módulos de elección están implicados en el proceso, a diferencia de lo que sucedería si fuéramos meros espectadores.

¿Podemos considerar que esta respuesta al «sofisma perezoso» es definitiva y concluyente?

Pienso que sí, si interpretamos el sofisma perezoso como un argumento en favor de una determinada forma de actuar. No hay razón para preferir la conclusión de «no hacer nada» a la de «ponerse en marcha». Dicho de otro modo, aceptar uno de los dos argumentos significa lo mismo, desde un punto de vista práctico, que admirar o querer parecerse a una persona cuyos módulos tengan una determinada configuración. Y en uno de ellos tal configuración sería el resultado de seguir el siguiente consejo: piensa en el futuro y en el curso del tiempo, y no hagas nada. Pero ¿por qué habría de admirar alguien a una persona que siguiera genuinamente este consejo? No sería más que un inútil: una persona que no haría tortillas ni alcanzaría cimas, y que ni siquiera lo intentaría.

Sin embargo, tal vez podemos interpretar de otra forma esta línea de pensamiento. Habitualmente, se considera que el fatalismo, más que recomendar un tipo de elección frente a otra, lo que hace es disolver cualquier posible elección. Pretende demostrar que nuestra capacidad de elegir no es más que una ilusión.

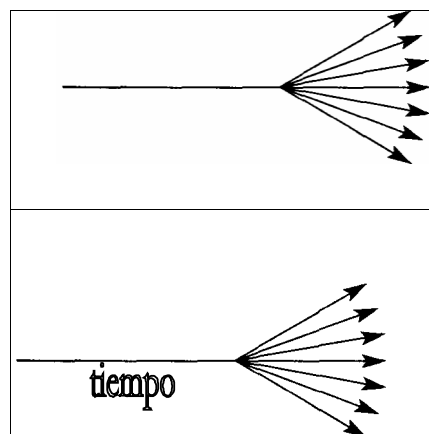
Pero ¿qué se supone que significa eso? Hemos defendido ya que cierta concepción de la elección es una ilusión. Se trata de la elección intervencionista, es decir, la intervención incausada y a gran escala del Verdadero Yo en el orden físico y neurofisiológico de los hechos. Nos hemos retirado a una reflexión acerca de los módulos flexibles de elección que se hallan implicados en nuestras acciones. ¿Cómo es posible que una serie de reflexiones acerca del paso del tiempo revelen que sus operaciones son irreales o ilusorias? Resulta tan poco plausible como decir que a causa del paso del tiempo el

funcionamiento de los ordenadores, de los termostatos y de las sierras mecánicas es ilusorio.

Cuando no sabemos lo que va a suceder, y creemos que nuestras acciones influirán en los acontecimientos, lo que hacemos es deliberar acerca de lo que podemos hacer. Hemos visto ya que el fatalismo no aporta ningún argumento que pueda orientar esta deliberación en un sentido o en otro. Y tampoco aporta ningún argumento que demuestre que el proceso es en sí mismo irreal, a menos que concibamos el proceso en los términos especulativos que antes hemos considerado y rechazado.

Sin embargo, supongamos que no sabemos lo que va a suceder, pero que alguien lo sabe, tal vez Dios. O simplemente que es algo cognoscible. Cuando deliberamos, consideramos que el futuro está abierto, mientras que el pasado es algo fijo. Pero supongamos que el futuro estuviera tan fijado como el pasado. En tal caso pensaríamos del siguiente modo:

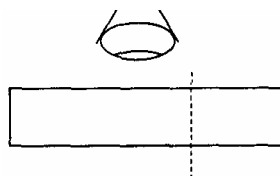
—donde las flechas representan



posibilidades abiertas, que se despliegan a partir del presente—. Pero tal vez esta forma de pensar sea ilusoria. Tal vez la verdad se vea sólo desde la perspectiva del «ojo de Dios», o lo que se ha llamado la «perspectiva desde el no tiempo». Desde esta perspectiva, el tiempo se despliega como una película de celuloide; un fotograma de la película corresponde a los sucesos de cualquier momento determinado. Tal como funciona el mundo, sólo somos conscientes de

los fotogramas del pasado (hay gente que cree que los profetas pueden «ver» los fotogramas del futuro). Pero no hay ninguna asimetría metafísica entre el pasado y el futuro:

«ojo de Dios»



presente en movimiento

Si eso fuera verdad, sería seguramente tan inútil tratar de influir sobre el futuro como tratar de hacerlo sobre el pasado. Si Dios dispone de esta perspectiva, debe de estar riéndose al contemplar nuestros esfuerzos. Eso es lo que se desprende de la parábola sufí. La Muerte ya tiene escrita su lista. Y por eso los soldados de mi amigo usaban la metáfora de una bala que «llevaba un número escrito», lo que implica que «ya llevaba un número escrito», es decir, con independencia de lo que pudieran hacer.

Pero ¿por qué habrían de reírse Dios o la Muerte? Supongamos que Dios dispone de una perspectiva intemporal. Sigue siendo cierto que no puede ver una tortilla en un determinado momento sin que nadie haya roto un huevo en un momento ligeramente anterior. Dios sabe si en un determinado fotograma de la película nos comeremos una tortilla. Pero en ese caso también sabe si en un fotograma ligeramente anterior nos pondremos a preparar la tortilla. No hay razón para que sepa cómo será el futuro con independencia de lo que hagamos, como tampoco puede saber que cierto árbol va a caer con independencia de lo que haga el viento. Desde la atalaya intemporal, todo lo que se ve es el viento y la destrucción. Dios no es, según todos los indicios, algo así como un médico que supiera que nos iba a matar un cáncer con independencia de lo que hiciéramos. Eso significaría que habría fotogramas en los que la gente se comportaría

de formas muy distintas y, sin embargo, moriría igualmente de cáncer. La «perspectiva desde el no tiempo», exterior al tiempo, permite ver nuestras acciones y sus resultados, pero no resultados sin acciones. Dios nos ve comer una tortilla porque nuestros módulos de elección nos pusieron a romper un huevo. Y sólo puede vernos comiendo una tortilla si en un fotograma anterior nos ha visto romper un huevo.

Lo que se desprende de la parábola sufí es que la Muerte tenía al discípulo en la lista antes de que éste decidiera huir, de modo que, según parece, le habría ido a buscar allí donde estuviera, fuese en Bagdad o en Samarcanda. Por este motivo, su huida era inútil. Pero quizá la Muerte lo tenía en su lista precisamente a causa de su huida, por ejemplo porque a su llegada a Samarcanda fuera a parar bajo las ruedas de un autobús. En tal caso, habría sido el hecho de huir lo que le habría llevado a su destino, aunque eso no nos dice nada acerca de si el discípulo se comportó de un modo razonable. Si la Muerte estaba haciendo su agosto en Bagdad, por ejemplo porque hubiera una epidemia en la ciudad, entonces la huida habría sido bastante racional, aunque desafortunada en el caso concreto. Podría haber ocurrido que la muerte no lo tuviera en su lista, precisamente a causa de su huida.

¿Qué podemos decir entonces acerca de la asimetría entre el pasado y el futuro? Si son simétricos a los ojos de Dios, ¿qué tiene de racional tratar de cambiar el futuro? ¿Cómo puede ser más racional que esforzarse por cambiar el pasado? Pues bien, tal como he dicho, ni siquiera Dios puede ver cómo nos disponemos a preparar una tortilla, si poco antes nos ha visto comerla, a menos que nos vea preparar y devorar ávidamente segundas tortillas. Así pues, es inútil tratar de influir sobre el pasado. Eso abre, sin embargo, un enorme e impracticable problema filosófico. Pues, ¿acaso el hecho de que no seamos capaces de influir sobre el pasado no podría ser una cuestión puramente fáctica, una contingencia que podría no haberse dado o

que podría ser de otro modo en regiones distintas del espacio y el tiempo? Si todo ello depende simplemente de las pautas que se advierten desde la perspectiva intemporal, parece que debería ser así. ¿Es posible que las pautas sean distintas en otros lugares?

Por el momento dejaré este problema como ejercicio para el lector (un ejercicio extremadamente difícil). Volviendo a la cuestión del fatalismo, hay que concluir que no tiene en realidad ninguna justificación general de tipo filosófico o racional. Corresponde a cierta actitud, un estado de ánimo en el que nos vemos impotentes y nos sentimos como meros espectadores de nuestras propias vidas. Eso no siempre resulta injustificado. A veces las personas se ven en buena medida impotentes, por motivos políticos o incluso psicológicos (porque no tengan la debida flexibilidad, a causa de un lavado de cerebro, o porque estén atrapados por extrañas obsesiones a las que no puedan sustraerse). En los momentos de impotencia, es natural que recaigamos en el fatalismo como actitud mental. Si nuestros mejores esfuerzos fracasan demasiado a menudo, necesitamos un consuelo, y la idea de un destino infinito en constante despliegue, un karma, puede ser a veces un consuelo.

Sin embargo, no es una actitud aconsejable cuando se trata de actuar. Cuando conducimos un coche, no es seguro pensar que es indiferente si giramos el volante o tocamos el freno. Nuestros esfuerzos no son indiferentes.

La flexibilidad y la dignidad

La ideología del dualismo mente-cuerpo está profundamente asentada. Cuando hablo de «ideología» no me refiero a un argumento específico o un conjunto de argumentos, sino más bien a un esquema de pensamiento: un punto de referencia o una idea que da sentido a todo lo demás. A menudo se pretende que el dualismo es lo único que hace posible la libertad, la dignidad y la propia experiencia humana.

Da sentido a las grandes palabras: el tipo de palabras que se escriben sobre las banderas. En los dos últimos capítulos he tratado de desvincular estas nociones. Pero la alternativa da miedo a mucha gente. ¿Acaso pretendemos reducir a las personas, en toda su animada y vistosa complejidad, a grises y aburridas máquinas monocromas, condicionadas para ser de una forma o de otra, o incluso peor, a meros vehículos pasivos de nuestros genes egoístas?

En absoluto.

El problema en este caso es que las alternativas se plantean como si agotaran todas las posibilidades: o bien tenemos un espíritu libre, que flota felizmente más allá del orden natural, o una máquina predeterminada como un autobús, o incluso un tranvía.

Plantear de forma errónea las alternativas es una conocida falacia, y la volveremos a encontrar en capítulos posteriores. Lo que denigra la naturaleza humana no es la filosofía compatibilista, sino este modo de formular las alternativas. Este planteamiento supone que la naturaleza es algo tan horrible que se requiere un momento mágico, una chispa divina surgida del fantasma de la máquina, para que muestre sus colores. Sólo puede haber máquinas (zombis) o espíritus. Pero es esta concepción lo que denigra la naturaleza, incluida la naturaleza humana. Deberíamos aprender a pensar como Wittgenstein cuando escribía:

Resulta humillante tener que aparecer como un tubo vacío, simplemente hinchado por una mente.⁷

La palabra clave que hay que tener en cuenta es «flexibilidad» (recordemos otra vez aquellos zombis inflexibles, programados). Resulta imposible determinar *a priori* lo flexible que puede ser el comportamiento humano. Digamos que nuestra biología es la responsable de nuestros módulos. Pero cómo vayan a funcionar estos módulos —cómo vayan a ser programados, si preferimos decirlo así, por su entorno— es otra cosa. Del mismo modo, la biología nos da

ciertas estructuras, sean lo que sean, pero igualmente tenemos que aprender a hablar. Poseemos esas estructuras; ningún otro animal las posee en un nivel remotamente parecido. Pero qué idioma aprendemos es algo que no viene determinado por la biología, sino por el entorno, ya que los niños imitan el habla de sus padres y demás personas cercanas.

Del mismo modo, nuestra capacidad de percibir cosas, de considerar las alternativas posibles, de evaluarlas, y finalmente nuestras pautas de comportamiento, podrían haber sido altamente inflexibles. Pero la evidencia sugiere que se da el caso contrario. Las personas aprenden naturalmente a valorar las cosas más diversas. Resulta bastante difícil descubrir siquiera alguna pauta universal: domina la flexibilidad. Los seres humanos son capaces de hacer campos de minas, pero también de hacer jardines.

A los teóricos y a los gurús les gusta establecer pautas: todo el mundo es egoísta; la gente sólo se mueve por intereses de clase; los hijos odian a sus padres; las personas son fácilmente influenciables; los hombres son agresivos; las mujeres son dóciles; la gente no sabe resolver sus problemas, y así sucesivamente. Pero eso no es tanto plegarse a la evidencia como imponerle una determinada interpretación. Como todos los estereotipos, tales interpretaciones resultan peligrosas, ya que las personas pueden acabar por adaptarse a ellas y convertirse en algo bastante más pobre de lo que habrían sido de otro modo. En este caso, el trabajo de ingeniería conceptual consiste en ofrecer un esquema más claro de otras posibles formas de pensar, y hay muchas.

CAPÍTULO

4

El yo

Ya hemos visto un aspecto del contenido de nuestra mente al tratar la cuestión de la conciencia. También nos hemos ocupado de nuestra actividad como sujetos y de nuestra libertad, es decir, de nuestra forma de relacionarnos con el mundo. Pero ¿qué podemos decir de nuestra propia subjetividad: el «yo»? Tal como vimos, Descartes no pudo rescatar nada más del naufragio de la duda universal. También vimos cómo Lichtenberg ponía en duda incluso que tuviera algún derecho a hacerlo. ¿Quién tenía razón y qué deberíamos pensar acerca del yo?

¿Un alma inmortal?

Veamos algunas de las cosas que pensamos acerca de nosotros mismos:

☐ Lista 1

Hubo un tiempo en el que yo era muy pequeño.

Salvo accidente o mala suerte, envejeceré.

A medida que envejezca, es probable que olvide buena parte de mis

recuerdos. También experimentaré algunos cambios, por ejemplo en relación con las cosas que me interesan. Mi cuerpo también experimentará cambios.

La materia orgánica de mi cuerpo (a excepción de mi cerebro) se renueva aproximadamente cada siete años.

Si mi cuerpo sufriera un accidente, por ejemplo la pérdida de alguna de sus partes, me tendría que conformar con el resultado.

Veamos ahora algunas cosas que es posible pensar acerca de nosotros mismos. Cuando digo que es posible pensarlas, me refiero a que parece que somos capaces de comprender lo que significan, no que creamos en ellas. Las posibilidades planteadas nos pueden resultar bastante peregrinas, pero eso no es relevante por el momento:

□ Lista 2

Podría haber nacido en una época o en un lugar distintos.

Podría ser que sobreviviera a mi muerte corporal, y viviera otra clase de vida como espíritu.

Podría haber merecido la dicha o la desgracia de tener un cuerpo distinto.

Podría haber merecido la dicha o la desgracia de tener unas capacidades mentales distintas (una mente distinta).

Podría haber merecido la dicha o la desgracia de tener a la vez un cuerpo y una mente distintos.

Podría ser que fuera la reencarnación de un personaje histórico.

Podría tener que vivir otra vez, por ejemplo en la forma de un perro, a menos que me comporte bien.

En realidad, hay gente que cree, o dice que cree, en tales cosas, y existen religiones enteras que sostienen algunas de ellas. El cristianismo sostiene que la segunda idea de esta lista es verdadera, mientras que el hinduismo defiende la última. E incluso aunque no aceptemos ninguna de ellas, parece que entendemos lo que

significan.

La diferencia entre ambas listas es la siguiente. La primera es compatible con la concepción que podemos tener espontáneamente sobre nosotros mismos. Soy un animal de gran tamaño, humano. Mi biografía es parecida a la de otros animales: comienza con un nacimiento natural, incluye una serie de cambios naturales y termina con una muerte natural. Estoy rigurosamente ubicado y limitado en el espacio y en el tiempo. Sobrevivo a varios cambios naturales, como el envejecimiento. Pero eso es todo.

La segunda lista sugiere que soy algo mucho más misterioso, algo que sólo de un modo contingente se halla «atado a un animal mortal». De acuerdo con las ideas que aparecen en la segunda lista, soy algo que puede cambiar de aspecto y de forma, de cuerpo y de mente, y que podría existir incluso sin ningún cuerpo. La biografía del «yo» podría durar siglos enteros y revelar una lista interminable de cambios de personalidad, como si fuera un actor.

Tal como vimos en los dos primeros capítulos, Descartes pensaba que podíamos intuir de forma «clara y distinta» que el yo es algo distinto del cuerpo. Y las ideas que se contemplan en la segunda lista parecen apoyarle. Es como si hubiera algo —el alma, el yo o la esencia— que permanece a través de un buen número de cambios (lista 1) y que incluso podría sobrellevar situaciones mucho más extraordinarias (lista 2). Pero, ¿qué es este yo? Veamos otra vez lo que dice David Hume:

En lo que a mí respecta, siempre que penetro más íntimamente en lo que llamo mí mismo tropiezo en todo momento con una u otra percepción particular, sea de calor o frío, de luz o sombra, de amor u odio, de dolor o placer. Nunca puedo atraparme a mí mismo, en ningún caso sin una percepción, y nunca puedo observar otra cosa que la percepción. Cuando mis percepciones son suprimidas durante un tiempo: en un sueño profundo, por ejemplo, durante todo ese tiempo no me doy cuenta de mí mismo, y puede decirse que verdaderamente no existo. Y si

todas mis percepciones fueran suprimidas por la mente y ya no pudiera pensar, sentir, ver, amar u odiar tras la descomposición de mi cuerpo, mi yo resultaría completamente aniquilado, de modo que no puedo concebir qué más falta para convertirme en una perfecta nada.¹

Hume está señalando que el yo es inaprensible. Es inobservable. Si «miras en tu interior» intentando atraparlo, no lo consigues porque todo lo que encuentras son lo que él llama «percepciones particulares», es decir, experiencias y emociones. No encuentras junto a ellas ni un atisbo del «yo» que es el sujeto de tales experiencias. Y sin embargo todos creemos conocernos a nosotros mismos de un modo particularmente íntimo. Tal como vimos, Descartes pensaba que este autoconocimiento sobrevivía incluso a la duda «hiperbólica».

A muchos filósofos les ha parecido que este diamante en bruto, el yo, poseía otra notable propiedad. Es simple. El yo no es un compuesto. Veamos lo que dice un contemporáneo de Hume, el filósofo escocés del «sentido común» Thomas Reid (1710-1796):

Hablar de una parte de una persona es manifiestamente absurdo. Cuando un hombre pierde su patrimonio, su salud, su fuerza, sigue siendo la misma persona, y no ha perdido nada de su personalidad. Si le cortan una pierna o un brazo, es la misma persona que era antes. El miembro amputado no era una parte de su persona, pues de otro modo tendría derecho a una parte de su patrimonio y sería responsable de una parte de sus obligaciones. Tendría derecho a una parte del mérito o el demérito que mereciera esta persona, lo cual es manifiestamente absurdo. Una persona es algo indivisible... Mis pensamientos, acciones y sentimientos cambian a cada momento; su existencia no es continua, sino sucesiva; pero el yo al cual pertenecen es permanente, y mantiene la misma relación con todos los pensamientos, acciones y sentimientos sucesivos que llamo míos.²

Este «yo» simple y perdurable es precisamente lo que Hume

lamentaba no poder encontrar nunca, Reid da un golpe sobre la mesa y anuncia su existencia.

La simplicidad del alma tiene la ventaja de abrir oportunamente la puerta al argumento tradicional sobre su inmortalidad.

Todo cambio o corrupción se debe a la conjunción o a la desintegración de cosas compuestas. *Luego*, cualquier cosa que no sea un compuesto no puede cambiar ni corromperse. El alma no es un compuesto. *Luego*, el alma no puede cambiar ni corromperse.

Tal como está formulada, la primera premisa podría resultar escasamente convincente. Haría falta dar algunas explicaciones. La idea sería que en cualquier cambio natural (físico) podemos detectar algo que se conserva. Si rompemos una galleta, la materia de la galleta se conserva. Antes se pensaba que se conservaban los átomos, de modo que los cambios químicos no serían más que una reordenación de los átomos de una sustancia. En la actualidad parece que deberíamos pensar en algo más básico todavía: tal vez lo que se conserva es la energía, o ciertas partículas subatómicas cuya reordenación sería la responsable de los cambios en la materia compuesta. En ambos casos, sólo cambiaría la composición. La verdadera «materia» (partículas elementales, energía) simplemente sigue allí.

Si pudiéramos defender realmente que la primera premisa es una verdad *a priori*, y si pensamos que Reid ha dado buenas razones en defensa de la segunda premisa (el alma no es un compuesto), entonces el argumento tiene buen aspecto. Por supuesto, el argumento también vale en relación con la existencia de mi «yo» antes de mi nacimiento natural, lo que podría hacer que perdiéramos un poco el entusiasmo.

¿Es posible que todas estas reflexiones sean ilusorias? ¿Debemos aceptar realmente que la lista 2 ofrece siquiera meras posibilidades? No nos preocupemos de momento por si aquellas posibilidades

existen realmente, tal como defienden numerosos creyentes. Preguntemos en cambio si por lo menos son coherentes.

Robles y barcos

Es útil reflexionar acerca de lo extrañas que resultan algunas de las creencias planteadas en la segunda lista. Elevan al yo por encima de todo cuanto parece darle alguna identidad, sea el cuerpo, la historia, la memoria o incluso la mente. ¿Tiene esto algún sentido? Para acercarnos a esta cuestión, comencemos por apartar la atención de nosotros mismos y examinemos la identidad de otros objetos. Podemos acudir de nuevo a John Locke, quien hizo una observación interesante a propósito de las plantas:

Siendo, pues, lo que constituye la unidad de la planta esa organización de sus partes en un cuerpo coherente que participa en una vida común, una planta continúa siendo la misma mientras siga participando de la misma vida, aun cuando esa vida sea comunicada a nuevas partículas de materia, unidas vitalmente a la planta viva, en virtud de una semejante organización continuada, que sea la conveniente a esa especie de planta.³

Locke señala el hecho de que un roble, por ejemplo, sigue siendo el mismo a lo largo de un determinado período de tiempo, a pesar de que cambien los «átomos» constituyentes (las células o las moléculas). La condición que se requiere es «participar de la misma vida» o, en otras palabras, lo que podríamos considerar una unidad organizativa o funcional. No importa si las partes son las mismas, mientras se mantenga esta unidad de función. Y en la medida en que sea así, podemos decir con propiedad que se trata del mismo roble, de modo que el roble es el mismo cuando es un primer brote, cuando es un árbol maduro, cuando se le caen algunas ramas, y así sucesivamente.

Esta intuición le sirve a Locke para explicar por qué identificamos al mismo ser humano a lo largo de los cambios normales de la vida. A estos efectos, no hay diferencia entre decir «el mismo hombre o mujer», «el mismo árbol» o «el mismo mono». El crecimiento y el cambio son integrados en la medida en que existe una continuidad de función o de vida organizada. Hasta aquí no hay problema. Locke tiene bien identificado aquello que nos permite reconocer al mismo ser humano (tomado como un mamífero de gran tamaño: lo que vemos cuando nos miramos en el espejo) o a la misma planta a lo largo del tiempo. ¿Por qué habría de ser distinto en el caso del yo?

Si examinamos las ideas de la segunda lista que he presentado al comienzo de este capítulo, nos daremos cuenta de que no podemos aplicarlas a los animales y a las plantas. No tienen ningún sentido. Nunca se nos ocurriría pensar, a propósito de cierto roble: «Mira, este árbol podría haber sido un arce», a menos que nos estemos refiriendo a que allí donde plantamos un roble podríamos haber plantado otro árbol, por ejemplo un arce. Pero en ese caso estaríamos hablando de un árbol distinto. No sería aquel mismo roble disfrazado, por así decirlo, de arce. Asimismo, no se nos ocurre pensar que los árboles sobrevivan a la muerte orgánica, de modo que aquel mismo árbol pudiera regresar, por ejemplo, en la forma de un narciso. Así pues, si no hay diferencia entre ser «el mismo yo» y ser «el mismo ser humano», y si establecemos la identidad de los seres humanos a través del tiempo de modo parecido a como la establecemos en el caso de los animales, entonces parece que ninguna de las ideas de la lista 2 debería tener sentido.

Un mismo roble no tiene por qué ser un agregado de idénticas moléculas en dos momentos distintos. Se puede decir lo mismo incluso de los objetos inorgánicos. Pensemos en una nube que resbala sobre la cima del Everest. Al montañista le parece que la misma nube tarda horas o días en alejarse de la cima. Y sin embargo está cambiando de composición a cada momento, a medida que el

viento desplaza las moléculas de agua a más de cien kilómetros por hora. A pesar de ello, sigue siendo la misma nube. Toleramos las diferencias de composición, por lo menos hasta cierto punto. Pensamos del mismo modo cuando nos referimos a grupos humanos, como clubes o equipos. Nos parece que animamos al mismo equipo año tras año, a pesar de que cambian los integrantes del equipo, y posiblemente su directiva y su campo. La gloriosa historia del regimiento no sería la mitad de gloriosa si no fuéramos capaces de identificar al regimiento más allá de sus miembros actuales. Pensamos del mismo modo también cuando se trata de objetos inanimados que poseen cierta función. Sigue siendo el mismo ordenador, por más que yo grabe otras cosas en su memoria, cambie la pantalla, actualice el sistema y otras cosas por el estilo.

A menudo nos importa bien poco si un objeto sufre más o menos cambios a la hora de seguir considerando que es el mismo «objeto»: valga como ejemplo aquel chiste sobre un hacha irlandesa que la familia conservó durante generaciones, aunque tuvo tres hojas y cinco mangos distintos. A veces se dan casos dudosos: buena ilustración de ello es el caso del «barco de Teseo». Teseo parte hacia un largo viaje, en el curso del cual varias partes de su barco deben ser reemplazadas. En realidad, hacia el final del viaje ha lanzado por la borda velas, mástiles, aparejos y tablas, y los ha reemplazado por otros nuevos. ¿Podemos decir que vuelve con el mismo barco? Probablemente nuestra respuesta sería que sí. Pero supongamos que algún oportunista le va siguiendo, recoge las piezas descartadas y las vuelve a ensamblar. ¿Acaso no puede reclamar que posee el barco original? Pero ¿acaso puede haber dos barcos distintos, cada uno de los cuales sea idéntico al original?

Almas y bolas elásticas

Así pues, tal vez deberíamos invocar algún tipo de «sustancia

inmaterial» (el alma, misteriosa y simple) para dar sentido a las ideas contenidas en la lista 2. Y si estas ideas nos parecen lo bastante sólidas, podríamos pensar incluso que constituyen un argumento en favor del dualismo cartesiano, a partir del hecho de que sólo tienen sentido en aquel contexto.

Pero en este punto Locke realiza una maniobra extraordinariamente interesante. Hemos visto ya que los animales y las plantas sobreviven al cambio de sustancia material. De modo que ¿por qué no habrían de sobrevivir las personas (yo, tú) al cambio de sustancia espiritual?

Pero la cuestión es averiguar si puede ser la misma persona cuando se cambia la sustancia misma que piensa, o si, permaneciendo ésta sin cambio, puedan ser personas diferentes.

A esto contesto, primero, que ésta no puede ser cuestión para todos aquellos que pongan el pensamiento en una constitución animal puramente material, privada de una sustancia inmaterial. Porque, independientemente de que su suposición sea o no verdadera, es llano que conciben que la identidad personal se conserva en algo que no es la identidad de la sustancia, a la manera en que la identidad animal se conserva en la identidad de vida y de sustancia. Y por lo tanto, quienes pongan el pensar en sólo una sustancia inmaterial tienen que mostrar, antes de atacar la otra opinión, por qué motivo la identidad personal no pueda conservarse en el cambio de sustancias inmateriales, o en una variedad de sustancias particulares inmateriales, de la manera en que la identidad animal se conserva en el cambio de sustancias materiales, o en una variedad de cuerpos particulares.⁴

La maravillosa maniobra de Locke pretende mostrar que por mucho que nos preocupe la supervivencia personal a través del tiempo y de los cambios, invocar «sustancias inmateriales» no representa ninguna ayuda. ¿Por qué no? Porque del mismo modo que reconocemos a las plantas a través del tiempo con independencia del cambio de los elementos materiales, también reconocemos a las

personas a través del tiempo sin necesidad de ninguna «sustancia inmaterial». Kant ofrece una buena ilustración de esta idea. En el siguiente pasaje de su obra maestra, la *Crítica de la razón pura*, las «representaciones» son cosas como experiencias o pensamientos (lo que Descartes habría agrupado bajo el nombre de *cogitationes*), es decir, el contenido de la mente:

Una bola elástica que choca con otra igual y que se mueve en la misma línea recta le comunica todo su movimiento y, consiguientemente, todo su estado (si sólo atendemos a las posiciones en el espacio). Supongamos ahora, por analogía con estos cuerpos, unas sustancias de las que una comunica sus representaciones y la conciencia de éstas a la otra. Surgiría así una serie de sustancias, la primera de las cuales transmitiría su estado y la conciencia de éste a la segunda; ésta, a su vez, su estado propio y el de la anterior a la tercera; ésta, igualmente, los estados de todas las sustancias precedentes, juntamente con la conciencia propia y la de las demás. La última sustancia sería, pues, consciente de todos los estados de las sustancias previamente cambiadas como de los estados propios, ya que tales estados, juntamente con la conciencia de los mismos, le habrían sido transmitidos.⁵

La idea es que no sabemos nada acerca de las «sustancias inmatrimales». Podría ser que nuestra sustancia inmaterial fuera reemplazada cada noche por una nueva, como si se tratara de un cambio de disco en un ordenador que conservara todo el *software* y los archivos.

Todas estas consideraciones bastan para lanzar serias dudas sobre el argumento de la inmortalidad que considerábamos antes. Tal como lo expresa Kant:

En efecto, no podemos apreciar nosotros mismos si, en cuanto almas, somos o no permanentes, ya que sólo asignamos a nuestro yo idéntico aquello de lo que tenemos conciencia. Por ello debemos, en cualquier

caso, juzgar necesariamente que somos los mismos en todo el tiempo del que poseemos conciencia. Ahora bien, no podemos dar todavía por válido este juicio desde el punto de vista de otro, ya que, al no encontrar en el alma otro fenómeno permanente que la representación del yo, la cual las acompaña y enlaza todas, nunca podemos establecer si este yo (mero pensamiento) no fluye igual que los otros pensamientos que él liga entre sí.

Podemos resumir su crítica diciendo que no hay nada en nuestras elucubraciones internas acerca de un «mí mismo» que permita pensar en una sustancia permanente, capaz de sobrevivir a los cambios y a las situaciones más extraordinarias. Pero, tanto Locke como Kant tienen ideas más constructivas que ofrecer.

El bravo oficial

Locke afirma que «es el hecho de tener una misma conciencia lo que hace que un hombre sea sí mismo para sí mismo» y no un sujeto ni terceras personas que garanticen que la conciencia es «poseída» por una sustancia permanente o por una sucesión de distintas sustancias permanentes. Locke lleva todavía más lejos el énfasis que ha puesto en la conciencia al afirmar que una persona A, en un determinado momento, es la misma que la persona B, anterior en el tiempo, sólo si A es consciente de las experiencias de B. En otras palabras, A debe recordar haber pensado lo mismo que pensó B, y recordar haber percibido, sentido y actuado del mismo modo que B percibió, sintió y actuó.

La sugerencia tiene algunas implicaciones que podrían resultarnos atractivas. Por ejemplo, elimina la posibilidad de que yo pueda ser la reencarnación de Cleopatra, ya que no soy consciente de haber hecho o sentido nada de lo que Cleopatra pudiera haber hecho o

sentido. El vacío de recuerdos destruye la identidad personal. De modo parecido, puedo estar seguro de que no viviré otra vida como perro, ya que ningún perro podría recordar haber hecho las cosas que he hecho yo; si las recordara (pensemos en la complejidad neuronal que se requiere) ya no sería un perro, sino, en el mejor de los casos, un ser humano con forma de perro. Pero los perros no son seres humanos con forma de perro.

Por otro lado, la sugerencia tiene algunas implicaciones que podrían resultar menos atractivas. Por ejemplo, no podría sobrevivir a una amnesia completa: sea quien sea la persona que permanezca después de tal circunstancia, no podría ser yo. Pero también tiene problemas con la amnesia parcial. Supongamos que cometo un crimen, pero que luego, quizás a consecuencia del inmoderado acceso de sangre o adrenalina, no conservo ningún recuerdo del episodio en cuestión. Parece seguirse de la teoría de Locke que no soy la persona que cometió el crimen. Soy el mismo ser humano, pero no la misma persona. Parece como si un único ser humano fuera habitado por múltiples personalidades sucesivas, en función del ir y venir de los recuerdos.

Thomas Reid planteó una versión de este problema en su «objeción del bravo oficial»:

Imaginemos a un bravo oficial que hubiera sido azotado en la escuela por haber robado en un huerto, que se hubiera apoderado de un estandarte enemigo en su primera campaña y que en su madurez hubiera sido nombrado general; supongamos también, ya que es algo que debemos admitir como posible, que cuando se apoderó del estandarte era consciente de haber sido azotado en la escuela, y que, cuando le nombraron general, era consciente de haberse apoderado del estandarte, pero había perdido toda conciencia de haber sido azotado. Partiendo de estos supuestos, se sigue de la doctrina del señor Locke que la persona que fue azotada en la escuela es la misma persona que se apoderó del estandarte, y que ésta es la misma persona que fue

nombrada general. De lo cual se deduce, si hay alguna verdad en la lógica, que el general es la misma persona que aquella que fue azotada en la escuela. Sin embargo, la conciencia del general no alcanza hasta sus tempranos azotes; por lo tanto, según la doctrina del señor Locke, no es la misma persona que fue azotada. Por lo tanto el general es, y al mismo tiempo no es, la misma persona que fue azotada en la escuela.⁶

En realidad, el propio Locke era perfectamente consciente de este problema. Su réplica es sencilla:

Sin embargo, es posible que se ofrezca una objeción. En efecto, supongamos que pierda por completo la memoria de algunas partes de mi vida, sin que haya posibilidad alguna de recobrarla, de manera que quizá yo no vuelva jamás a tener conciencia de esos actos. Ahora bien, ¿no soy, acaso, la misma persona que llevó a cabo esos actos, que tuvo esos pensamientos, de los cuales alguna vez tuve conciencia, aunque ahora los he olvidado? A lo cual responde que aquí es necesario advertir a qué se aplica la palabra yo; o sea, que en este caso, designa solamente al hombre. Y como se supone que el mismo hombre es la misma persona, se supone sin dificultad que la palabra yo también significa «la misma persona». Sin embargo, si es posible que un mismo hombre tenga, en diferentes momentos, distintas e incommunicables conciencias, no hay duda alguna de que un mismo hombre sería diferentes personas en diferentes momentos.⁷

Podemos reconstruir el argumento de la siguiente forma. O bien la expresión «la misma persona» corresponde a «el mismo ser humano», o bien no corresponde. En el primer caso, todos estaremos de acuerdo en que el mismo ser humano se mantiene desde la infancia hasta la muerte, con independencia de sus capacidades mentales. Por lo tanto, ninguna de las ideas de la lista 2 tiene el menor sentido. Si Locke niega que «la misma persona» corresponda a «el mismo ser humano», es porque si consideramos que un ser humano puede tener «distintas e incommunicables conciencias», tendremos que pensar que hay diferentes personas habitando

sucesivamente un mismo cuerpo (podemos pensar en cierto tipo de desórdenes mentales, como la personalidad múltiple). Pero en tal caso la sorprendente afirmación de que el viejo general no es la misma persona que el escolar resulta ser correcta.

La concepción de Locke se basa, en cierto sentido, en la misma idea que Reid se negaba a aceptar. Según éste, lo que necesitamos es una noción de «la misma persona» a través del tiempo que nos permita justificar las imputaciones de responsabilidad. Reid pensaba que la identidad de la persona era una noción jurídica, es decir, que su espacio natural eran los tribunales de justicia. La cuestión queda más clara si consideramos el caso de un tembloroso octogenario que fuera inesperadamente acusado de crímenes cometidos, digamos, durante una guerra que hubiera tenido lugar sesenta y cinco años antes, cuando él no era más que un ingenuo recluta adolescente. ¿Es eso justo? Supongamos que nuestro hombre responde con toda sinceridad que no conserva ningún recuerdo de sus crímenes. En tal caso, desde su punto de vista sería como si le condenaran por unos actos cometidos por una persona completamente distinta. Y eso parece injusto: si una persona no es consciente de sus actos, tampoco puede «arrepentirse» de ellos, simplemente porque no forman parte de su propia autoconciencia. No pueden pesarle sobre la conciencia.

Por supuesto, Locke era consciente de que nuestros tribunales no trabajan de esta manera. Después de todo, la amnesia no es una excusa válida. Sin embargo, pensaba que eso no respondía más que a una cierta suspicacia nuestra, debida al hecho de que resulta muy fácil alegar amnesia. A los ojos de Dios, una verdadera amnesia sí valdría como excusa. Dios trataría al sincero amnésico octogenario como si fuera una persona distinta de aquella que una vez fue un criminal de guerra. Esta idea podría resultar atractiva, aunque no tanto en el caso del crimen cometido por culpa de un acceso de sangre a la cabeza; nosotros diríamos que su autor no lo ha olvidado, ni en este mundo ni en el otro. Nos interesaría señalar distintos

grados en la pérdida de memoria.

¿Tiene razón Reid cuando acusa a la teoría de Locke de contravenir la propia lógica y llevar a una contradicción? Este tipo de contradicción se conoce como «error en la transitividad» de la identidad. La transitividad es la regla lógica según la cual si $A = B$ y $B = C$, entonces $A = C$. En este caso, el escolar = el oficial, y el oficial = el general, pero según Locke no es cierto que el escolar = el general. A esto se refiere Reid cuando habla de contradicción.

Visto así, el argumento de Locke resulta ciertamente extraño, pero tal vez la extrañeza procede del hecho de que hablemos de «identidad» en términos abstractos, cuando en realidad nos referimos a «ser la misma persona que». Consideremos una vez más un compuesto cualquiera, como por ejemplo una bicicleta o un barco. Supongamos que la antigüedad de un barco es relevante, por ejemplo a efectos de la tarifa fiscal que se le deba aplicar. Podría ser que los barcos antiguos pagaran menos impuestos. ¿Cuándo se puede decir que un barco es genuinamente antiguo? Imaginemos que Teseo y el oportunista que recogió las piezas originales de su barco reclaman cada uno por su lado la rebaja de impuestos. Si estos fastidiosos oportunistas se convirtieran en algo habitual, tal vez deberíamos aprobar una ley que estableciera cuál de los barcos es el original. La ley podría decir algo parecido a esto:

Todo barco debe ser registrado anualmente, y para que sea considerado el mismo barco en relación con cualquier año anterior, el navío deberá contener al menos el 55 % de los materiales que lo integraban el primer día después de su registro de aquel año.

Ahora apliquemos el esquema de Reid: se ve fácilmente que bajo este código, $Argos_1$ sería el mismo barco que $Argos_2$, y $Argos_2$ el mismo barco que $Argos_3$, pero $Argos_1$ no sería el mismo barco que $Argos_3$. A pesar de ello, la ley resulta bastante razonable en su

formulación actual, muy en la línea de las leyes que especifican lo que deben contener ciertos productos para que sean considerados mantequilla o pienso. Y nadie pensaría que una ley razonable pudiera llevar a una contradicción.

Pues bien, resulta que los barcos son objetos compuestos, hechos de partes, y parece que eso es lo que origina el problema. De modo que el argumento de Reid, según el cual no es posible que $A=B$, $B=C$, pero no $A=C$, tal vez sólo sea válido en caso de que tanto A como B y C sean simples, y no compuestos. En este caso, tal como vimos, Reid sostuvo que el alma era simple, mientras que Locke no, de modo que quizás el argumento no vale contra él.

El yo como manojó de experiencias

Tal como antes señalaba Hume, cuando reflexionamos acerca del contenido de nuestra propia mente encontramos recuerdos, pensamientos, pasiones y experiencias particulares, pero ningún yo. Hume sostiene que no tenemos derecho a hablar de aquello que no encontramos (ni podemos encontrar) en la experiencia. Nuestra mente no puede aprehenderlo, ni siquiera «rozarlo». Coherentemente, sostiene que el yo no es más que un agregado de nuestras «percepciones» o experiencias y de las conexiones que pueda haber entre ellas. Habría un contenido, pero no un continente. Esto es lo que a veces se llama teoría «no posesiva» del yo, o teoría del «manojó». Al igual que Lichtenberg en el primer capítulo, Hume sostiene que hay un «se piensa», o quizás un «pensamiento en curso». Pero no hay un propietario o poseedor, o un «yo» al que atribuir este pensamiento.

El problema habitual con este planteamiento es que exige darle algún sentido a la idea de una experiencia sin propietario. Sin embargo, se objeta, eso sería incoherente. Significaría tratar las experiencias como «objetos» o cosas por derecho propio: el tipo de

cosa que podría estar flotando libremente, sin propietario, a la espera de ser atrapado en un manojo junto a otras experiencias, como los palos en el bosque. Sin embargo, prosigue la objeción, pensar eso es un error, puesto que las experiencias son parasitarias o adjetivas en relación con las personas que las poseen. ¿Qué sentido tiene esto?

Consideremos las abolladuras de los coches. Podemos hablar sobre abolladuras: esta abolladura es peor que aquella otra, o será más cara de reparar que la del año pasado. Pero es lógicamente imposible que haya una abolladura «sin propietario», una abolladura sin una superficie abollada. La abolladura es, por así decirlo, un producto derivado de un adjetivo. Al comienzo tenemos una superficie, la cual experimenta un cambio al ser abollada, y luego abstraemos un sustantivo y nos ponemos a hablar de la abolladura. El sustantivo «abolladura» depende lógicamente del adjetivo, «abollada». De modo parecido, una sonrisa depende de una cara que sonrío, y con esa idea jugaba Lewis Carroll cuando escribió que el gato Cheshire desapareció dejando sólo la sonrisa.

Así pues, a Hume se le objeta que las «experiencias» son parasitarias de las personas en el mismo sentido. Nos resulta imposible imaginar un dolor, por ejemplo, como una «cosa» que flotara libremente, en espera de ser atrapada en un manojo de otras experiencias, de modo que resultara accidental el hecho de que ese dolor en concreto terminara en un manojo o en otro. En un principio está la persona, y el dolor comienza cuando una parte de la persona comienza a doler, del mismo modo que una abolladura comienza a existir en el momento en que una superficie se abolla.

Kant plantea esta misma idea en relación con el «yo pienso» que acompaña a todas mis representaciones. En otras palabras, mis experiencias vienen marcadas como «mías». No es que primero tenga conocimiento de una experiencia, luego busque a su propietario, y finalmente (en caso de que la búsqueda resulte fructífera, contra la opinión de Hume) anuncie que la experiencia es

mía. Al contrario, decir que percibo un dolor es lo mismo que decir que algo me duele.

Pero, ¿cómo es eso posible, si Hume está en lo cierto cuando dice que nunca percibimos un «yo»? Está muy bien que comparemos los dolores y las abolladuras, y es cierto sin duda que si percibo una abolladura es porque percibo una superficie abollada. Pero al menos se puede decir que percibo una superficie, abollada o no. En cambio, si Hume está en lo cierto no parece que percibamos un alma o un yo.

Tal vez la forma de evitar del problema sea negar que el «yo» sea el tipo de cosa que se puede percibir. Wittgenstein examinó aquellos casos en los que nos describimos a nosotros mismos como sujetos de la experiencia: «Oigo la lluvia» o «Tengo dolor de muelas». Su conclusión fue que en estos casos «no hay problema a la hora de identificar a la persona». «Tan imposible es que al decir "tengo dolor de muelas" me esté confundiendo con otra persona, como lo es que me queje de dolor por equivocación, tomándome por otra persona.» No puedes equivocarte al identificarte a ti mismo como el sujeto de una experiencia. Wittgenstein cree que eso lleva a una ilusión:

Vemos, pues, que en los casos en los que se usa «yo» como sujeto, no lo usamos porque reconozcamos a una persona particular por sus características corporales; y esto crea la ilusión de que nosotros utilizamos esta palabra para referirnos a algo incorpóreo que, sin embargo, tiene su sede en nuestro cuerpo. De hecho, éste parece ser el ego real, aquel del que se dijo *Cogito, ergo sum*. «¿No hay entonces mente, sino solamente un cuerpo?» Respuesta: la palabra «mente» tiene significado, es decir, tiene un uso en nuestro lenguaje; pero el decir esto no dice todavía qué tipo de uso hacemos de ella.⁸

Deberíamos tratar de concebir la autoconciencia de un modo distinto. ¿Cuál?

El yo como principio organizativo

Traslademos el problema al campo de la inteligencia artificial. Imaginemos un robot equipado con una videocámara que es capaz de desplazarse por una habitación en la que se han dispuesto una serie de objetos. Supongamos que nuestro objetivo es conseguir que el robot nos entregue una descripción de la disposición de los objetos en la habitación. ¿Qué deberíamos hacer? Cuando el robot dirige su cámara hacia un objeto, los píxeles se activan. Así pues, tiene la clase de «brillo interior» que a veces se asocia a la conciencia. Pero si eso es todo lo que percibe, no es más que una «rapsodia de percepciones», según la expresión de Kant, o una «florecente y susurrante confusión», según la expresión que empleó más tarde el psicólogo estadounidense William James (1842-1910). En otras palabras, el robot todavía tiene que organizar los datos antes de poder interpretar la escena. Supongamos que la pantalla muestra una pequeña forma redonda. ¿Se halla cerca de un objeto pequeño o lejos de uno grande? ¿Acaso está viendo un objeto elíptico de lado? Para resolver estos problemas, el robot debería moverse para obtener una imagen distinta y luego realizar una «síntesis» de las diversas imágenes, antes de poder construir una representación tridimensional de la habitación. ¿Qué habilidades se requieren para realizar esta síntesis? ¿Cómo se pueden unificar una serie de imágenes distintas obtenidas en momentos distintos?

Parece que los ingredientes básicos son los siguientes. El robot necesita saber de alguna forma si él mismo se está moviendo. En concreto, necesita ser capaz de distinguir si se está moviendo, y obtiene por lo tanto nuevas imágenes de objetos en reposo, o si está quieto, y son los objetos que le rodean los que se mueven. Para ello necesita tener un recuerdo de cómo era antes la habitación para poder compararlo con su aspecto actual. Necesita ser capaz de reconocer el orden de las diferentes imágenes, y luego necesita ser capaz de integrar de alguna forma las imágenes anteriores con la imagen actual. En otras palabras, para resolver el problema de la

posición que ocupan los objetos en el espacio, necesita resolver el problema de cuál es su propio punto de vista y saber cuánto tiempo ha transcurrido, para poder así cartografiar sus propios movimientos.

Lo que sugiere este ejemplo es que un mínimo de autoconciencia es un requisito estructural para cualquier posible interpretación de la experiencia. El programador no puede resolver el problema en el caso del robot simplemente mediante la incorporación de otro ingrediente en la pantalla (por ejemplo, que la cámara recogiera siempre un fragmento de una de las ruedas del robot, que se podría ver en la parte inferior de la pantalla). Eso no sería más que otro *input*. Se requiere otro tipo de programación para traducir los *inputs* en una descripción de la habitación y del lugar que ocupa el robot en ella.

En realidad, el robot no necesita ninguna imagen de sí mismo. La cámara puede estar rígidamente orientada hacia la escena que tiene delante. Por esta razón, si Hume hubiera percibido un elemento constante en la experiencia, como una nota de fondo, al explorar su propio interior, no habría estado por ello más cerca de aprehenderse a sí mismo. Lo que necesita el robot, en cambio, es encontrar el modo de representar su propia ruta en el espacio y de establecer el orden temporal de las imágenes que obtiene. Es un requisito esencial para la solución del problema que posea un punto de vista «egocéntrico» o, en otras palabras, que elabore una representación del espacio centrada en «él mismo». Si fuera capaz de interpretar que tiene una mesa a tres pies de distancia, también tendría que ser capaz de decir que «la mesa está a tres pies de mí» y, sin embargo, no tendría por qué saber nada de su propia forma corporal, o de su pasado lejano. Y con toda seguridad no necesitaría estar familiarizado con su propio ego interno o su alma inmortal.

Si hubiera el suficiente caos en la habitación, el problema podría convertirse en irresoluble. Por ejemplo, si tuviéramos la mala idea de poner al robot en un entorno parecido al de los Keystone Cops, donde los objetos vienen y van al azar o a una velocidad vertiginosa, se

vería atrapado en un problema insalvable: los píxeles se activarían al azar, pero sin la necesaria continuidad entre un momento y el siguiente como para que algún programa pudiera encontrarle un sentido.

Así pues, pensar en términos de un «yo» parece ser una condición necesaria desde el punto de vista formal o estructural para que sea posible interpretar la experiencia tal como lo hacemos nosotros: como experiencia de un mundo tridimensional de objetos permanentes, entre los cuales nos movemos. El «yo» es el punto de vista que hace posible la interpretación. No es un elemento más de la experiencia, ya que no hay nada en la experiencia que pueda resolver el problema formal para el que se requiere el «yo». Al contrario, siempre se necesita un punto de vista: representarse una escena implica representarse a uno mismo experimentándola de una u otra forma.⁹

Engaños de la imaginación

La línea de pensamiento que acabo de exponer se la debemos a Immanuel Kant. Se trata de uno de los grandes argumentos que se han planteado en el campo de la filosofía, y sus repercusiones se han dejado sentir en todos los terrenos; más adelante volveremos sobre algunas de estas repercusiones. Sin embargo, para nuestros propósitos actuales su interés reside en el diagnóstico que sugiere acerca de las ideas contenidas en la lista 2, que aparece al comienzo de este capítulo.

Lo que hace que surjan aquellas ideas es mi supuesta capacidad para imaginarme a mí mismo en diferentes situaciones, y para ponerme en la piel, por ejemplo, de un personaje histórico, de un perro o de un ángel. Eso me hace pensar que he trasladado al escenario imaginado a mi misterioso yo, a mi propia alma, en cuyo caso el alma se convierte en algo realmente extraño, pues algunas de

mis fantasías pueden consistir en imaginarme a mí mismo en otra época, con un cuerpo distinto o con capacidades mentales distintas, viviendo otro tipo de experiencias y otras cosas por el estilo. En otras palabras, me abstraigo de todo aquello que me confiere una identidad como ser humano y, sin embargo, sigo pensando que hay algo, mi esencia, que permanece. Este es el origen de la «distinción real» de Descartes.

Pero supongamos, en cambio, que no traslado nada con mi imaginación. Todo lo que hago es representarme cómo sería el mundo visto desde otro punto de vista, en otra época, o lo que fuera. Si no hay ninguna esencia mía que se traslade a los diversos escenarios, el hecho de que sea capaz de imaginarlos no constituye ninguna evidencia de que mi «yo» los haya experimentado o pueda sobrevivir para experimentarlos. A modo de ilustración, consideremos una de las primeras ideas de la lista: tal vez sobreviva a mi muerte corporal. ¿Qué fantasías se esconden detrás de esa idea? Pues bien, supongamos que me imagino lo que sería contemplar mi propio funeral, en presencia de mi ataúd y de mi familia que se lamenta. Tal vez los vea desde un escondite situado al fondo de la iglesia. Tal vez me ofenda el hecho de que la congregación no parezca muy apenada, o quizá tenga ganas de decirles que a fin de cuentas no es tan terrible lo que me ha pasado. Tal vez el hecho de estar muerto me dé una visión de rayos X, de modo que echo un vistazo a mi cuerpo tendido dentro del ataúd. Se me encoge el corazón. Parezco tan viejo. Pero ¡mira! Allí se abren las puertas de nácar y allí viene mi bisabuela a darme la bienvenida...

Imaginar todo esto significa ensayar ante mí mismo la experiencia de ver mi propio ataúd y demás. Y eso es algo que sin duda puedo hacer: después de todo puedo representarme lo que sería verlo — algo no muy distinto de ver otros ataúdes—. Puedo representarme lo que sería echar un vistazo a su interior —una visión espantosa—. Sin embargo, y éste es el punto crucial, tales ejercicios de imaginación no

trasladan a ningún «yo» al escenario donde el ser humano Simon Blackburn está muerto. Soy yo, aquí y ahora, quien se lo imagina, y no hay ningún yo imaginario contemplando aquella escena. El único vestigio mío que hay en ella es mi cuerpo muerto.

Podemos plantear la cuestión del siguiente modo. La línea de pensamiento de Kant sugiere que es equivalente decir «puedo imaginar lo que es ver X» y «puedo imaginarme a mí mismo viendo X». Pero se trata de una equivalencia puramente formal: no hay ningún yo sustantivo, ningún alma o Mí Mismo, implicado en la fantasía. Por lo tanto, es un error tomar aquellas fantasías como fundamento de ninguna «distinción real» entre uno mismo como sujeto, como yo o como alma, y el animal que somos en realidad. Así pues, nuestras fantasías acerca de X no suponen ningún apoyo para la teoría de que nuestra biografía podría superar la biografía de aquel animal, simplemente porque X es algo que el animal no podrá ver.

De modo parecido, supongamos que practico un juego que podríamos llamar «imaginar que soy Genghis Khan». Me represento caballeros y campos de batalla. Soy bajo y astuto, y también un magnífico jinete. Por Dios, qué frías son las estepas. Tanta política acaba por aburrirme. Grito: «Otra ronda de leche fermentada de yegua». Uy, se supone que hablo mongol, no inglés.

En este caso debería resultar todavía más evidente que no hay ninguna alma ni Mí Mismo que se haya introducido en la figura de Genghis. En realidad, desde el momento en que sigue habiendo algo de mí en la fantasía, como por ejemplo el lapsus con el inglés, el intento ha sido un fracaso. Sucede lo mismo cuando un actor representa el papel de un personaje histórico y cae en anacronismos, como por ejemplo que Enrique VIII se mire el reloj o hable acerca de las películas que hay en cartelera.

Lo que realmente hago es representarme campos de batalla, frías estepas y otras cosas por el estilo, como si los tuviera delante de los ojos, y hago cosas dignas de un guerrero, como dar instrucciones o

saltar sobre un caballo. Puedo hacerlo más o menos bien: algunas personas tienen más facilidad para imaginarse el mundo visto desde diferentes puntos de vista, del mismo modo que hay personas que actúan mejor que otras. Mientras mi Genghis Khan siga hablando en inglés, sin embargo, no habré llegado demasiado lejos.

¿Demuestra eso que todas las ideas de la lista 2 son ilusiones? Ciertamente pone en duda el valor que puedan tener las simples fantasías como prueba. Si hay alguna otra evidencia en su favor, pues perfecto. Pero es bueno darse cuenta de lo mucho que debe aquella lista a las fantasías en primera persona. No parece que me lleve muy lejos tratar de imaginarte a ti en la persona de Genghis Khan. ¿Acaso te puedo imaginar masacrando a gente desde un caballo? ¿Con un género distinto, una edad distinta y una forma también distinta de pensar (se hace difícil creer que puedas tener las mismas ideas que Genghis Khan)? Todo lo que consigo cuando intento imaginar esta posibilidad es pensar en ti en lugar de pensar en Genghis Khan. Es lo mismo que pensar en un narciso en lugar de pensar en un roble, lo cual no es lo mismo que pensar que el roble podría haber sido un narciso. No consigo descubrir ninguna clase de identidad entre ambos.

En resumen, no puedo evitar pensar en ti como un animal humano de gran tamaño con una determinada personalidad. Otros animales humanos con otras personalidades son distintos de ti, y no podrías haber sido uno de ellos. ¿Cuántos aspectos de tu personalidad podrías perder y seguir siendo tú mismo? La verdad es que este problema se parece al que teníamos antes con los barcos. Tal vez estemos dispuestos a aceptar muchos cambios y, sin embargo, a veces decimos cosas como: «Ya no es la persona que era antes». Desde la perspectiva que proponen Locke y Kant, la expresión podría ser cierta en un sentido literal.

Recomponer el alma

Cuando pensamos sobre nosotros mismos, descubrimos que existe una curiosa diferencia entre el pasado y el futuro.

Supongamos que vivimos en un mundo en el que es más fácil montar y desmontar los cuerpos y los cerebros. Podemos separarlos en partes y volverlos a montar tal como hacemos con los ordenadores o los coches. Supongamos que tales operaciones se llaman «operaciones de recomposición». Al terminar podemos volver a cargar la psicología de las personas, de modo parecido a como copiamos *software* y archivos en un ordenador. O bien podemos cambiar la configuración, conservando parte del *software* y los archivos anteriores, y añadiendo otros nuevos. Las operaciones de recomposición son vistas como algo saludable y beneficioso.

Supongamos que en este mundo te dicen que al día siguiente te van a someter a una operación de recomposición. También te permiten saber algunas cosas acerca del resultado de la operación. La persona A conservará muchos de tus rasgos y muchas de tus cualidades: él o ella recordará las mismas cosas que tú, tendrá un aspecto físico muy parecido al tuyo y otras cosas por el estilo. En cualquier caso, la persona A será enviada al Ártico (puede que seas un miembro del ejército). La persona B también se parecerá a ti e incorporará también muchos de tus rasgos físicos actuales —el cerebro y las células—, así como tus cualidades (*software* y archivos). La persona B será enviada a los trópicos.

Desde nuestro punto de vista, el problema se parece mucho al del barco de Teseo. No tenemos por qué convertir en un gran dilema la cuestión de si te has convertido en la persona A o en la persona B. Podríamos considerar que eres una de las nuevas personas, o incluso ambas, o bien podríamos considerar que son creaciones nuevas. El filósofo contemporáneo David Lewis sugiere la analogía de un camino que se bifurca. No consideramos que sea un gran dilema metafísico la cuestión de si optamos por decir que sólo una de las rutas es la antigua carretera de Turnpike, o si ambas lo son, o si no lo es

ninguna.

Pero desde tu punto de vista, la cuestión parece ser crucial. O bien pasas frío durante un año, o bien pasas calor, o bien ni siquiera sobrevives. Sólo hay tres opciones tajantes. No puedes perderte en vaguedades e indeterminaciones: no tiene sentido decir: «Será un poco como si estuvieras en los trópicos y un poco como si estuvieras en el Ártico». En último término, no hay nadie que experimente algún tipo de mezcla entre los trópicos y el Ártico, entre el calor y el frío. A tiene frío, y B no. No puede ser mitad y mitad. Del mismo modo, decir que «será un poco como si no existieras y un poco como si existieras» tiene tan poco sentido como lo anterior. O bien estarás sudando en aquel preciso lugar, o bien te estarás congelando en el otro, o bien te habrás reunido con tus antepasados. «Estarás allí en la medida en que ellos estén allí» suena a camelo, como si alguien me consolara por no haber ido nunca a Venecia diciéndome: «Irás cuando vaya tu hijo».

Mejor que se calle. (Tal como dijo Woody Allen a propósito de un consuelo parecido: «No quiero alcanzar la inmortalidad a través de mi obra. Quiero alcanzar la inmortalidad a través de no morir».)

Lo extraño es que cuando pensamos acerca del pasado perdemos esta sensación de ruptura tajante. Supongamos que en aquel mundo recibes la noticia de que eres el resultado de una operación de recomposición en la que participaron dos personas, C y D, las cuales aportaron tal o cual rasgo a la persona que eres actualmente. La noticia resulta interesante, pero no te despierta aquella necesidad urgente y desesperada de saber. Si te dicen que C pasó las Navidades de 1990 en un barco y que D las pasó en la cima de una montaña, pero tú no recuerdas ni lo uno ni lo otro, la cuestión no tiene por qué obsesionarte: «¿Dónde estaba yo el día de Navidad de 1990?». Si la recomposición te dio una vaga conciencia de ambas experiencias, tampoco habrá problema: es un poco como si aquel día hubieras subido una montaña y un poco como si hubieras ido a navegar.

Resulta escalofriante darse cuenta de que en último término no habrá nadie que tenga problemas con su identidad. La persona A, que está en el Ártico, se halla en una relación de parcial continuidad contigo, y lo mismo sucede con la persona B, que está en los trópicos. Cualquiera de los dos puede recordar con nostalgia algunas de tus acciones. Y si les apetece pueden desear tal o cual parte de tu cuerpo, o bien tal o cual rasgo psicológico o recuerdo tuyo, del mismo modo que podemos contemplar con nostalgia nuestras personalidades pasadas y desear parecemos a ellas en mayor o menor medida. Podemos lamentarnos de las capacidades y los recuerdos perdidos, o bien alegrarnos del conocimiento y la madurez alcanzadas, según las preferencias de cada uno.

Hay gente que cree que se pueden encontrar soluciones concretas acerca del futuro. Puede que algunos confíen en la posibilidad de que su identidad sobreviva mientras su cerebro actual se mantenga activo. Por supuesto, Locke rechazaba esta posibilidad, ya que la continuidad del funcionamiento cerebral no asegura en lo más mínimo la continuidad de la conciencia: el cerebro podría ser «reprogramado» o reconfigurado hasta alterar completamente los recuerdos y la personalidad. Y en cualquier caso podemos imaginar alguna operación de recomposición que desmonte el cerebro y distribuya las partes. Puede que otros confíen en una continuidad de *software* (tipo Locke) antes que en una continuidad del *hardware*. Pero se enfrentan al problema de que en un mundo en el que fuera posible recomponer a las personas tal vez se podría copiar el *software* a placer y crear un buen número de futuras personas con idénticos «recuerdos» y rasgos de personalidad.

En resumen, no parece haber ninguna correspondencia metafísica entre la simplicidad del futuro tal como lo imaginamos nosotros y las complejidades e indeterminaciones que plantea la hipótesis de la recomposición.

Algunos pensadores pierden la paciencia ante esta clase de

escenarios. Afirman que nuestra noción de identidad está hecha a la medida del mundo real, en el que, tal vez por fortuna, las operaciones de recomposición son imposibles. Pretenden que dejemos en paz los problemas de la identidad en estos casos imaginarios y rebuscados. Mi opinión personal es que se equivocan. Estoy de acuerdo con ellos en que deberíamos dar menos importancia al problema de la identidad cuando se plantean hipótesis rebuscadas. Pero no creo que debamos menoscabar la importancia de la tarea de pensar acerca de nosotros mismos: acerca del hecho de que las opciones que se nos plantean son de naturaleza tajante, por muchas que sean las vaguedades que se esconden detrás de nuestros rasgos y componentes como animales. Sospecho que esto es lo que anima las reflexiones de muchas personas acerca de la vida y la muerte. Da pie a la esperanza y a la fe. Da pie a que algunas personas decidan congelar su cerebro, con la esperanza de que algún día, cuando la tecnología lo permita, serán descongelados y comenzarán una nueva vida, siendo ellos mismos. Da pie a la creencia de Reid de que el alma es simple. Un alma simple, que no pueda ser dividida, es un requisito indispensable para mantener abiertas aquellas tres opciones tajantes. Irá a parar a un sitio o al otro.

Sin embargo, es posible que nuestro apego por las opciones tajantes se base en una ilusión: la misma clase de ilusión que encontrábamos en las fantasías de la sección anterior. Entonces insistimos en que no había ningún «yo» que se trasladara a los escenarios imaginados. Ahora deberíamos insistir en que no podemos trasladar ningún yo definido a aquellos escenarios futuros. A medida que la evidencia acerca de qué animal humano estará presente en aquellos escenarios se vuelve vaga e indeterminada, la evidencia acerca de quién será la persona que estará presente en ellos se vuelve también vaga e indeterminada. Nuestra tendencia a pensar de otro modo se basa en una ilusión. Quizás nos resultaría más fácil descubrirla si recordáramos la razón por la cual Hume no podía dar

con su propio «yo», y la razón por la cual la explicación kantiana de la necesidad de pensar en términos de algún tipo de sujeto no nos ofrece más que un fundamento formal. Un diamante en bruto o un átomo de mí mismo, por más simple que sea, es incapaz de cumplir la tarea que el yo tiene encomendada.

Sin embargo, creo que puedo asegurarle al lector que le costará deshacerse de aquellas tres opciones tajantes. La reflexión puede ser una ayuda, aunque no resulta fácil destruir las ilusiones del yo.

Así pues, la «distinción real» que Descartes creía haber demostrado —el dualismo cartesiano— no da su brazo a torcer con facilidad. El lector es libre de intentar protegerla de la línea de argumentación que he seguido en este capítulo y en los dos anteriores. Valga decir que el propio Kant trató de dejar abierta la posibilidad de la inmortalidad del alma.¹⁰ Su razonamiento, más bien débil, es que necesitamos pensar que la bondad trae consigo la felicidad y, como en esta vida eso no siempre sucede y ni siquiera es razonable confiar en ello, sería mejor que hubiera otra vida en la que se cumpliera. Es entonces cuando las personas reciben lo que se merecen. En opinión de la mayoría de filósofos, no es éste uno de los mejores momentos de Kant. Pero el aspecto religioso de la cuestión ciertamente influye en la manera de pensar de muchas personas acerca de este tema, de modo que quizá será mejor dirigir nuestra atención hacia eso.

CAPÍTULO

5

Dios

Para algunas personas, pensar acerca del alma es casi lo mismo que pensar acerca de la religión. Y pensar acerca de la religión es una de las ocupaciones más importantes de la vida. Para otras personas, es casi lo mismo que perder el tiempo. En este capítulo presento algunos de los argumentos que rodean esta cuestión. Al menos los argumentos no son una pérdida de tiempo, puesto que introducen algunos principios importantes acerca del pensamiento mismo.

Las creencias y otras cosas

Damos por supuesto que nuestras creencias son verdaderas. Viene a ser lo mismo decir «Creo que p » o decir «Creo que es verdad que p ». Nadie diría: «Creo que existen las hadas, pero no pienso que sea verdad que existan». Y parece que las personas religiosas creen en algunas cosas en las que otras personas no creen.

Sin embargo, no resulta inmediatamente evidente que la verdad sea relevante cuando se trata de la religión, o que las actitudes religiosas se comprendan mejor en términos de verdad y falsedad. Pues cabe la posibilidad de que la religión no sea una cuestión de

creencias, y que aquellas actitudes no sean propiamente creencias. Aceptar una determinada religión tal vez se parezca más a recrearse en un poema o aficionarse al fútbol. La cuestión podría ser más bien participar en un determinado conjunto de prácticas. Tal vez el sentido de estas prácticas sea solamente emotivo, o social. Tal vez los rituales religiosos no sirvan más que a determinados fines necesarios desde el punto de vista psicológico y social. Tal es la función de los ritos relacionados con el nacimiento o con la madurez, o la de los funerales. No tiene sentido preguntar si una ceremonia de matrimonio es verdadera o falsa. La gente no va a un funeral para escuchar verdades, sino para mostrar su dolor, o para comenzar a superar su dolor, o para meditar acerca de la vida que ha terminado. Resulta tan inadecuado preguntar si lo que se dice en ella es verdadero como preguntar si la *Oda a una urna griega* de Keats es verdadera. El poema funciona o no en un sentido bastante diferente, y lo mismo cabe decir de la catedral de Chartres o de una estatua de Buda. Puede que sean fastuosos, emotivos o solemnes, pero no porque impliquen afirmaciones que puedan ser verdaderas o falsas.

Algunos piensan que eso es todo cuanto se puede decir acerca de la cuestión. De modo que si alguien dice: «Dios existe» no es como si hubiera dicho: «El abominable monstruo de las nieves existe» (donde la cuestión es empírica) o «Existen números primos entre el 20 y el 30» (una cuestión matemática). Es más bien una forma de expresar alegría o miedo (o, en un sentido más siniestro, de expresar odio hacia los extraños o los infieles). Éste es el motivo de que sus afirmaciones sean inmunes a una crítica enfocada hacia su verdad o falsedad. En el mejor de los casos, podríamos examinar los estados mentales que tienen como resultado y tratar de evaluar si resultan admirables o no.

Sin embargo, no ha sido ésta la forma más común de interpretar la religión. Aunque admitiendo su aspecto emocional y social, muchos han pensado que estaban realizando afirmaciones concretas acerca

del mundo: afirmaciones verdaderas en un sentido literal, para las que habría pruebas y argumentos. Desde esta perspectiva, la creencia religiosa es como las demás creencias: un intento de describir cómo es el mundo, las cosas que contiene y la razón de los sucesos que tienen lugar en él. Desde esta perspectiva, un funeral no es verdadero ni falso, pero algunas de las cosas que se dicen durante la ceremonia sí lo son, como por ejemplo que volveremos a levantarnos de entre los muertos. Desde esta perspectiva, cuando alguien dice sinceramente que resucitará, no significa que haya escogido una forma metafórica, poética o emocionalmente cargada de decir cualquier otra cosa o de introducir algo de brillo en el mundo ordinario. Está anunciando una determinada expectativa suya, tan literal como la expectativa de hacer un viaje o como la expectativa de la llegada de un amigo.

En el presente capítulo discutiré una serie de creencias religiosas en términos de argumentos, fundamentos y pruebas. Doy por sentado que se pretenden verdaderas y, por lo tanto, que se ajustan a nuestras mejores estrategias para alcanzar la verdad. La mayoría de filósofos sólo se ha interesado por ellas desde esta perspectiva, aunque algunos filósofos morales, en especial Friedrich Nietzsche, han criticado las emociones y actitudes morales (humildad, sumisión y compasión) que según ellos fomentan ciertas religiones.

Para precipitar un poco las cosas, voy a presentar a continuación un cierto número de argumentos contrarios a la posibilidad de que alguna creencia identificable como religiosa sea verdadera. Esto puede hacer que algunos lectores se sientan atacados. En tal caso, pueden buscar consuelo en la tradición teológica que afirma que cuanto más improbable es una creencia, mayor es el mérito del acto de fe que se exige para creer en ella. Sin embargo, al final de este capítulo el inquieto espíritu de la reflexión nos llevará también a examinar ese planteamiento. En cualquier caso, comenzaré por examinar los argumentos filosóficos clásicos en favor de la existencia

de Dios: el argumento ontológico, el argumento cosmológico, el argumento del diseño y argumentos relativos a la revelación y a los milagros. Terminaremos por reflexionar acerca de la naturaleza de la fe, de la creencia y del compromiso.

El argumento de san Anselmo: bombones y pollos

Se cuenta la historia de un gurú que congregó en un estadio a una importante audiencia con la promesa de que ofrecería una prueba definitiva de la existencia de Dios. Cuando todos se hubieron reunido, exhibió con amplios ademanes el *Oxford English Dictionary* y mostró que contenía la palabra «Dios». Puesto que la palabra estaba allí, y había una definición, tenía que existir algo que correspondiera a ella. No tengo la menor idea de cómo pudieron sentirse aquellos que se encontraban entre el público, o si alguno de ellos llegó a pensar que el diccionario también mencionaba a Santa Claus y a las hadas, aunque los calificara expresamente de míticos o imaginarios. Pero resulta interesante reflexionar acerca del hecho de que puedan existir palabras con significado a las que no corresponda ningún objeto.

El motivo es que una cosa es definir un concepto y otra muy distinta que exista algún objeto que corresponda a la definición. Si tuvieras la intención de anunciarte en las páginas de contactos, podrías definir lo que buscas en tu pareja:

Persona reflexiva busca hincha futbolístico amante de la diversión, vegetariano y que toque el banjo. Abstenerse fumadores.

Esto define a tu compañero o compañera ideal: nos permitiremos llamarle Bombón. Pero podría darse la lamentable circunstancia de que no existiera ningún hincha futbolístico amante de la diversión, vegetariano, no fumador y que tocara el banjo. Tú decides lo que incluyes en la descripción, pero es el mundo quien decide si hay alguien que se ajuste a ella. Podría ser que Bombón no existiera.

La descripción resulta perfectamente inteligible. Define una serie de condiciones que en principio podrían corresponder a alguna persona. Se da el caso, simplemente, de que nadie corresponde a ellas. Una posible forma de expresarlo es decir que aquellos términos tienen un sentido, pero ninguna referencia. Sabes lo que quieres decir, pero no sabes si hay algo que corresponda a ello. No puedes inferir la referencia a partir del sentido, porque la cuestión de si existe una referencia depende de cómo sea el mundo, y no se puede resolver mediante el estudio o la consulta de un diccionario.

Puede que te fastidie descubrir que nadie corresponde a tu descripción. Pero tal vez se te podría ocurrir un plan para resolver el problema. ¿Por qué no introduces una posdata especificando que Bombón debe existir? Ahora el anuncio rezaría lo siguiente:

Persona reflexiva busca hinchable futbolístico, etc., *que exista*.

En este momento podrías pensar que has resuelto el problema por definición.

Pues bien, ciertamente es verdad que nadie te llamará para decir que cumple con todas las condiciones excepto la última. Sin embargo, cualquier persona que hubiera llamado después del primer anuncio también hubiera existido: «Llamo, luego existo» es una inferencia tan válida como «Pienso, luego existo». Y el hecho de que introdujeras la cláusula final no añade ni un ápice a las posibilidades que tenías de que alguien cumpliera las condiciones de partida, de modo que has malgastado el dinero con las dos últimas palabras. Decir «que exista» no añade ninguna nueva especificación al compañero ideal y no mejora las posibilidades de que exista realmente. En ocasiones, los filósofos se refieren a esta circunstancia diciendo que «la existencia no es un predicado», lo cual significa que añadir «Que exista» no es como añadir «Y que le guste la cerveza Guinness». Tú eres quien domina el sentido: puedes añadir lo que quieras a la descripción. Pero

es el mundo el que domina la referencia: establece si hay algo que cumple con tus condiciones.

Si esto último ha quedado claro, podemos pasar a los argumentos. En el capítulo 1 ya encontramos un argumento en favor de la existencia de Dios: el argumento de la marca de Descartes. No parecía demasiado convincente, y el propio Descartes introdujo un nuevo argumento para reforzarlo más adelante, en la quinta *Meditación*. Se trataba de una versión de un argumento mucho más antiguo, el argumento ontológico de san Anselmo (1033-1109). Según éste, Dios se definía por ser «más grandioso que ningún otro ser que podamos concebir». En el siguiente fragmento, san Anselmo se dirige «al insensato» que afirmaba de corazón que Dios no existía:

Pero cuando este mismo insensato me oye decir «algo más grandioso que cualquier otro ser que podamos concebir», seguramente entiende lo que le digo, y todo lo que entiende existe en su entendimiento, aunque no entienda que exista (en la realidad) [...] Así pues, incluso el insensato debe admitir que existe, al menos en su entendimiento, algo más grandioso que cualquier otro ser que podamos concebir, ya que lo entiende cuando se lo dicen, y todo aquello que se entiende existe en el entendimiento. Y sin embargo, aquello que es más grandioso que cualquier otro ser que podamos concebir no puede existir únicamente en el entendimiento, porque si existiera únicamente en el entendimiento, podríamos concebirlo como existente en la realidad, lo cual significaría mayor grandeza [...] [P]or lo tanto, no hay duda de que existe, tanto en el entendimiento como en la realidad, algo más grandioso que cualquier otro ser que podamos concebir.¹

Lo interesante acerca de este argumento es que es puramente *a priori*. Pretende haber demostrado la existencia de Dios a partir de un simple examen del concepto o definición de Dios. Se parece a cierto tipo de prueba matemática que deduce, a partir del concepto de circunferencia, el hecho de que las cuerdas trazadas desde un punto cualquiera hacia los extremos opuestos de un diámetro forman un

ángulo recto. El argumento no necesita premisas empíricas: ninguna medición o dato procedente de la experiencia.

El argumento de san Anselmo se puede presentar en dos pasos:

Entendemos el concepto de Dios. Cualquier cosa que entendemos existe en el entendimiento. *Luego* Dios existe en el entendimiento.

Y luego:

Supongamos que Dios sólo existiera en el entendimiento, y no en la realidad. En tal caso sería concebible un ser más grandioso que Dios: un ser que existiera en la realidad. Sin embargo, Dios se define como el ser más grandioso que se pueda concebir. *Luego*, por definición, no se puede concebir ningún ser más grandioso. Pero ahora nos encontramos ante una contradicción. *Luego*, nuestra hipótesis original era falsa.

Este tipo de argumento, que describo con mayor detalle en el próximo capítulo, se llama *reductio ad absurdum*. San Anselmo nos ha hecho asumir una premisa ateísta, pero sólo como paso previo para poder demostrar su falsedad, a partir del hecho de que implica una contradicción.

La versión cartesiana del argumento emplea el concepto de «perfección» más que el de grandiosidad, pero su estructura es parecida. Dios se define como perfecto, y el hecho de que no existiera sería, en cierto sentido, una imperfección en relación con sus otras cualidades: «La existencia es una perfección». Así, la existencia pertenece a la esencia de Dios, y no es posible concebir a Dios como no existente.

Ya en la época de san Anselmo, un fraile llamado Gaunilo cuestionó el argumento. Gaunilo observó que si el argumento fuera válido, también podría ser empleado para demostrar toda clase de conclusiones demasiado buenas para ser ciertas: por ejemplo, que existe una isla más perfecta que ninguna otra que se pueda concebir. Volviendo a nuestro Bombón, podemos aplicarle el mismo principio.

Supongamos que añades aplicadamente a las especificaciones de Bombón que no sólo debe ser un gran amante, sino el mejor amante que se pueda imaginar. En tal caso se podría argumentar de una forma parecida:

Entendemos el concepto de Bombón. Cualquier cosa que entendemos existe en el entendimiento. *Luego* Bombón existe en el entendimiento.

Y luego:

Supongamos que Bombón sólo existiera en el entendimiento, y no en la realidad. En tal caso sería concebible un amante mejor que Bombón: un amante que existiera en la realidad. Sin embargo, Bombón se define como el mejor amante que se pueda concebir. *Luego*, por definición, no se puede concebir ningún amante mejor que Bombón. Pero ahora nos encontramos ante una contradicción. *Luego*, nuestra hipótesis original era falsa.

Bombón existe en la realidad. ¡Fantástico! Pero no saltes de alegría demasiado pronto. Lamentablemente, podrías demostrar por los mismos medios que existe un rival tan peligroso como se pueda imaginar compitiendo contigo por el amor de Bombón. La premisa crucial en este caso sería que los rivales reales son más peligrosos que los meramente imaginarios, lo cual seguramente es cierto. El argumento ontológico parece capaz de demostrar también la existencia del Diablo, definido como lo más malo que se pueda concebir. Porque si algo hubiera de ser lo más malo que pudiera concebirse, mejor sería que no existiera sólo en la imaginación, pues en tal caso se podría concebir algo todavía más malo, a saber, un ser que fuera igual de malo pero que además existiera en la realidad (cabe señalar que en el caso de un diablo la existencia es una imperfección: lo hace todavía más malo).

La mayoría de filósofos han reconocido que hay algo sospechoso en el argumento ontológico (tan sospechoso como pretender que

Bombón se convierta en realidad gracias a una simple descripción). Pero no siempre han estado de acuerdo sobre dónde hay que buscar exactamente el error. Parte del problema consiste en la estrategia de considerar «la existencia como un predicado». Este problema se resuelve gracias a una teoría que trataremos en el próximo capítulo, la teoría de los cuantificadores. Pero se hace difícil asegurar que el error fatal consiste en esto.

Desde mi punto de vista, el problema de fondo hay que buscarlo en la ambigüedad que se esconde detrás de la comparación entre la «realidad» y el «entendimiento». El argumento compara las cosas que están «en la realidad» con las cosas que están «en el entendimiento» (es decir, en función de una definición, o bien en la imaginación o en los sueños), en relación con propiedades como la grandiosidad o la perfección. Parece sencillo, como si se tratara de comparar cosas que estuvieran en diferentes regiones geográficas, en cuyo caso sabríamos perfectamente si las que están en una región son más o menos grandes que las que están en la otra. Sería como preguntar si los pollos pesan más en Alemania que en Francia. Pero en realidad no tiene nada que ver con eso. Examinemos la siguiente oración:

Los pollos reales pesan más que los pollos imaginarios.

Parece que al menos en un sentido la oración es verdadera. En este sentido, se puede decir que los pollos imaginarios no pesan nada (después de todo, no conseguiríamos llenar ni un mísero plato con uno de ellos). Pero en otro sentido la oración resulta ser falsa, ya que podemos imaginar un pollo más pesado que ninguno que exista en la realidad: un pollo que pesara quinientos kilos y del tamaño de un pequeño granero, por ejemplo. En el caso del argumento ontológico, «Dios» en el entendimiento es comparado con Dios en la realidad, igual que hacíamos con el pollo imaginario y el pollo real, y se

concluye que el primero pesa menos. En el argumento anterior, el Bombón real es comparado con el Bombón imaginario, y se le considera mejor: ¡incluso un amante real bastante mediocre es mejor que uno imaginario! Y se supone que ello contradice la definición. Pero en realidad esta clase de comparación no revela nada en absoluto que contradiga la definición.

Es como si una profesora nos pidiera que imagináramos un pollo más pesado que ningún otro pollo existente. Supongamos que lo hacemos: imaginamos un pollo de quinientos kilos. Pero entonces la profesora nos reprocha que no hemos imaginado lo que ella nos pedía, puesto que los pollos imaginarios pesan siempre menos que los reales. El pollo que hemos imaginado no pesa nada (no nos lo podemos comer) y por lo tanto hemos «contradicho la definición» y no obtenemos ninguna calificación. En tal caso tendríamos derecho a sentirnos injustamente tratados. No somos nosotros los que nos hemos equivocado, sino la profesora.

Todo esto sugiere que no debemos pensar en los «pollos imaginarios» o en los «pollos en el entendimiento» como si fueran clases de pollos que en principio pudieran ponerse en una báscula junto a los pollos reales, aunque siempre resultaran pesar menos. Sin embargo, el argumento ontológico exige precisamente este tipo de comparaciones. Éste es su fallo. Porque incluso en el caso de que Dios no existiera más que en la imaginación, como sucede en el caso de Bombón o en el del pollo de quinientos kilos, no se sigue de ello que podamos describir o imaginar algún ser superior. Después de todo, la descripción incluía todos los superlativos. Pero por desgracia para la demostración de san Anselmo, eso no resuelve la cuestión de si existe alguna cosa que corresponda a ella.

Elefantes y tortugas

El argumento ontológico ha resultado siempre sospechoso. El

mejor teólogo y filósofo de la Edad Media, santo Tomás de Aquino (c. 1225-1274), se negó a aceptarlo. Prefirió defender que Dios es necesario para explicar el mundo o el cosmos tal como lo conocemos. Este argumento, conocido como el argumento *cosmológico*, resulta mucho más atractivo para la imaginación. Existen varias versiones de él. En todas ellas se trata de identificar un criterio según el cual todas las cosas que existen en el universo físico, las cosas que conocemos a través del tacto, de la vista y de los demás sentidos, son entes dependientes. Se argumenta entonces que los entes dependientes presuponen, en último término, como fundamento o explicación, un ente que no sea él mismo dependiente de ninguna otra cosa. Una de las versiones de este planteamiento, y quizá la más fácil de comprender, es el argumento de la causa primera. Veamos a continuación lo que dice el personaje de Demea en los *Diálogos sobre la religión natural* de Hume (estos *Diálogos*, publicados un año después de la muerte de su autor, en 1776, constituyen el análisis filosófico clásico de los argumentos teológicos tradicionales, y en las páginas siguientes aparecerán citados con frecuencia):

Todo lo que existe tiene necesariamente una causa o razón de su existencia, porque es absolutamente imposible que algo se produzca por sí mismo, o que sea la causa de su propio existir. Al remontarnos, pues, de los efectos a las causas, o continuamos en una sucesión infinita sin llegar a una última causa, o nos vemos obligados a recurrir, por fin, a alguna causa última, que *necesariamente* existe. Ahora bien, que la primera suposición es absurda se prueba del siguiente modo. En la cadena o sucesión infinita de las causas y efectos, cada efecto individual se dice que existe por el poder y eficacia de la causa que antecede inmediatamente; pero a la cadena o sucesión eterna, tomada en conjunto, no hay nada que la determine o cause y, sin embargo, es evidente que requiere una causa o razón, tanto como cualquier objeto particular que tiene su principio de existencia en el tiempo. Sigue siendo cuestión razonable saber por qué es precisamente esta sucesión

particular de causas la que existe desde la eternidad, y no otra sucesión, o ninguna. Si no hay un ser necesariamente existente, toda suposición que pueda hacerse es igualmente posible; y es igualmente absurdo pensar que nada haya existido desde la eternidad, que pensar en esa sucesión de causas, que constituye el universo. ¿Qué fue, pues, lo que determinó que algo existiese en lugar de nada, y confiriéndose el ser a una posibilidad particular, con exclusión de las demás? *Causas externas*, se supone que no las hay; el azar es una palabra que carece de significado. ¿Fue, acaso, nada? Pero eso jamás puede producir algo. Debemos, por lo tanto, recurrir a un Ser necesariamente existente, que lleve en sí la razón de su existencia, y que no pueda suponerse no existente sin contradicción expresa. En consecuencia, hay un tal Ser, es decir, hay una Deidad.²

El argumento resulta persuasivo, pero ¿es válido?

Fue Russell quien, supuestamente, descubrió que el argumento de la causa primera tenía un fallo y, por cierto, un fallo único y espectacular, que consistía en que no sólo la conclusión no se deducía de las premisas, sino en que de hecho las contradecía. Su idea es que el argumento parte de la premisa de que «todo tiene una causa previa y distinta» y, sin embargo, termina con la conclusión de que es necesario que haya algo que no tenga una causa previa y distinta, sino que «posea en sí mismo la razón de su existencia», en cuyo caso la conclusión niega lo que afirma la premisa.

La crítica de Russell resulta algo apresurada, ya que desde el punto de vista teológico el sentido del argumento es que aunque todo aquello que es material o físico tiene una causa previa y distinta, este mismo hecho nos lleva a postular alguna otra cosa que no la tiene. En la jerga teológica, esto último sería algo «necesario» o *causa sui*: algo que es causa de sí mismo. Y como esto no es cierto en relación con los objetos ordinarios que nos rodean, debemos postular algo extraordinario, una Deidad, que posea esta extraordinaria autosuficiencia.

En los *Diálogos* de Hume, el problema que tiene este planteamiento queda resumido en pocas líneas:

Se pretende que la Deidad es un ser necesariamente existente, y se trata de explicar esa necesidad de su existencia afirmando que si conociéramos en su integridad su esencia o naturaleza, nos daríamos cuenta de que le es tan imposible no existir como que dos veces dos no sean cuatro. Mas es evidente que, mientras nuestras facultades permanezcan en el estado presente, eso jamás ocurrirá. En cambio, en todo tiempo nos será posible concebir la inexistencia de aquello que con anterioridad concebimos como existente, y jamás puede la mente ser constreñida por la necesidad de suponer que algún objeto siempre permanezca siendo; de la misma manera como es constreñida por la necesidad de concebir siempre que dos veces dos son cuatro. Así pues, las palabras «existencia necesaria» carecen de sentido, o, lo que es lo mismo, no tienen ninguno que sea congruente.³

EL portavoz de Hume en este fragmento, un personaje llamado Cleantes, dice más adelante que el mundo o el universo material en conjunto, por lo que sabemos, debe ser también un ente de existencia necesaria, por mucho que sus partes dependan unas de otras. La razón es que sólo unas «cualidades desconocidas, inconcebibles», podrían hacer que algo fuera «necesariamente existente». Y por lo que sabemos, tales cualidades desconocidas e inconcebibles corresponderían antes al universo físico conocido que a ninguna cosa, persona o deidad inmaterial que se encontrara más allá de él.

Es importante recordar en este punto que toda nuestra experiencia cotidiana nos obliga a pensar que las mentes están tan necesitadas de explicación y son tan dependientes como los objetos físicos. Postular una mente que sea inmune de algún modo a cualquier tipo de dependencia significa tanto apartarse de la experiencia como postular un objeto físico que también lo sea.

El argumento de la causa primera responde a una inquietud

natural y, según algunos filósofos, en especial Kant, inevitable. Cuando pensamos en el *big bang*, la primera pregunta que se nos ocurre es: ¿y eso por qué sucedió? No nos satisface la respuesta «por ningún motivo», porque no nos satisface que las cosas sucedan «porque sí»: nuestro afán de obtener explicaciones es insaciable, de modo que postulamos alguna otra cosa, otra causa que se esconde detrás de ésta. Pero en este caso las explicaciones amenazan con multiplicarse al infinito. Ahora que hemos introducido a Dios, nos vemos obligados a preguntar cuál es la causa de Dios, o bien a cortar arbitrariamente la regresión. Pero si ejercemos este derecho arbitrario a detener la regresión en este punto, también podríamos haberla detenido en el universo físico. En otras palabras, nos encontramos en la misma posición que aquel filósofo hindú al que preguntaron qué era lo que sostenía al mundo y respondió: «Un elefante», y cuando le preguntaron qué era lo que sostenía al elefante, respondió: «Una tortuga», y cuando le preguntaron qué era lo que sostenía a la tortuga, pidió que cambiaran de tema.

Por otro lado, hay versiones del argumento cosmológico que no tienen en cuenta la cuestión de la causa primera. Se interesan más bien por el orden inmutable del universo: la uniformidad de la naturaleza. Puede parecer un hecho asombroso que las leyes de la naturaleza se mantengan, que la estructura de la naturaleza no se desmorone. Cabría pensar que estos hechos tienen que ser «dependientes» y que requieren una causa necesaria que los sostenga (igual que Atlas sostenía el mundo). Pero una vez más, o bien nos encontramos ante una regresión, o bien ante un mero decreto que afirma que algo posee «propiedades desconocidas e inconcebibles» que lo convierten en autosuficiente. La uniformidad imperturbable de esto último no requeriría ninguna explicación más allá de sí mismo. Y el mundo en conjunto es tan buen candidato como cualquier otra cosa. Pero la uniformidad de la naturaleza es algo que trataremos en los dos próximos capítulos.

El sabio arquitecto

El mismo Cleantes que se encarga de refutar el argumento cosmológico es el portavoz de una tentativa distinta de demostrar la existencia de una deidad: el argumento del «diseño» —la idea de que el cielo y la tierra manifiestan la gloria de su creador—. Éste fue el argumento estrella de la teología del siglo XVIII, y sigue ejerciendo una poderosa influencia. En lo que sigue voy a ceñirme a la discusión clásica que aparece en los *Diálogos* de Hume. Cleantes presenta el argumento:

Echad una mirada en torno al mundo; contemplad el todo y cada una de sus partes, veréis que no es otra cosa sino una gran máquina, subdividida en un infinito número de máquinas más pequeñas que a su vez admiten subdivisiones hasta un grado que va más allá de lo que los sentidos y facultades humanas pueden rastrear y explicar. Todas estas máquinas y hasta sus partes más nimias se ajustan entre sí, con una precisión que arrebató la admiración de todos los que las han contemplado. La singular adaptación de los medios a los fines de la Naturaleza entera, se asemeja exactamente, aunque en mucho excede, a los productos del ingenio humano, a los de los designios del hombre, de sus pensamientos, su sabiduría y su inteligencia. Si, por lo tanto, los efectos se asemejan entre sí, estamos obligados a inferir, por todas las reglas de la analogía, que también las causas son semejantes, y que el Autor de la Naturaleza se parece en algo a la mente humana, aun cuando sus facultades sean mucho más considerables en proporción a la grandeza de la obra que ha ejecutado. Por este, y sólo por este argumento *a posteriori*, podemos probar al mismo tiempo la existencia de una Deidad y su semejanza con la mente e inteligencia humanas.⁴

Es importante señalar dos cuestiones en relación con este argumento. En primer lugar, se trata de un argumento por analogía. El mundo se parece a los objetos diseñados por los seres humanos. Por lo tanto, del mismo modo que cuando encontramos un reloj es razonable suponer que su autor es humano, cuando descubrimos la

majestuosa estructura de la naturaleza también es razonable postular que su autor es divino. En segundo lugar, se trata de un argumento a *posteriori*, es decir, que parte de la experiencia, o de aquello que conocemos acerca del mundo al entrar en contacto con él. Y es en este punto donde se desinfla la evidencia en favor del diseño.

Cuando el darwinismo estuvo en condiciones de ofrecer una explicación natural de los procesos de adaptación que tienen lugar entre los complejos sistemas biológicos, el argumento comenzó a perder algo de lustre. Sin embargo, Hume (y también Kant) ya hizo las observaciones pertinentes sin necesidad de apoyarse en ninguna explicación alternativa sobre cuestiones tales como la adaptación biológica, lo cual tiene exactamente el mismo valor, ya que el argumento no se refiere principalmente a la biología, que sólo habla de una instancia particular dentro de los ajustes que se producen en la naturaleza. La cosmología abarca otros. Por ejemplo, de acuerdo con una estimación actual que cuenta con una amplia aceptación, el hecho de que las diversas constantes cosmológicas se ajusten unas a otras de modo que la vida organizada sea posible en algún punto del universo tiene una probabilidad de 1 contra 10 elevado a 10^{125} (una cifra inimaginable), de modo que quizá sí que hizo falta un sabio arquitecto que las ajustara.

Así pues, ¿cuál es el argumento que emplea Filo, el portavoz de Hume en los *Diálogos*, para tratar de refutar el argumento del diseñador? Filo señala que el argumento toma una de las operaciones que encontramos en la naturaleza, el pensamiento, y la convierte en una «regla para el todo».

Pero concedido que se pudiera tomar las maneras de operar de una parte de la naturaleza como cimiento de nuestro juicio sobre el origen del todo (lo que es inadmisibile), sin embargo ¿por qué elegir un tan pequeño, tan endeble, tan limitado principio como lo son la razón y los designios de los animales que habitan en este planeta? ¿Qué privilegio especial tiene esa pequeña agitación del cerebro que llamamos el

«pensamiento», para convertirla en modelo del universo entero? Ciertamente a eso nos inclina la parcialidad que tenemos hacia nosotros mismos; pero la sana filosofía debe guardarse cuidadosamente contra tan natural ilusión.⁵

Un argumento por analogía debe cumplir ciertas condiciones para ser fiable. Primero, que los elementos en los que se basa la analogía sean extremadamente parecidos. Segundo, que las explicaciones que se ofrezcan como probables estén basadas en la experiencia, lo cual significa que deberíamos disponer de tanta información como fuera posible acerca del tipo de causa que produce el efecto en cuestión. Por ejemplo, un agujero en un árbol es bastante parecido a un agujero en un cuerpo humano. Pero llevamos demasiado lejos nuestros razonamientos si suponemos «por analogía» que como es probable que el hombre muera a causa de uno, también lo es que muera el árbol a causa del otro. Necesitamos realizar más observaciones, tener una comprensión más elaborada de cómo suceden estas cosas antes de que sea razonable adelantar alguna inferencia de este estilo. Es precisamente esta clase de experiencia la que se halla lamentablemente ausente en el campo de la teología, pues no tenemos la menor idea de qué tipo de «cosas» son las que dan origen a universos físicos enteros.

Por otro lado, los parecidos son fáciles de encontrar, y Filo se divierte de lo lindo inventando algunos. En primer lugar, por mucho que se diga que el universo se parece a un reloj, lo cierto es que se parece todavía más a un vegetal:

Claramente se ve que el mundo se asemeja más a un animal o un vegetal, que a un reloj o un telar. Su causa, pues, más probablemente, se asemejará a la causa de aquéllos. La causa de aquéllos es la generación o vegetación; la causa, pues, del mundo, así podemos inferirlo, es algo semejante o análogo a la generación o a la vegetación.⁶

Por supuesto, un teísta replicará inmediatamente que esto no nos

lleva a ninguna parte, pues lo único que conseguimos es otra causa de tipo vegetal, cuyo origen habría que investigar de nuevo. Pero cabe decir lo mismo si vamos a parar a una causa del estilo de una mente. Si Cleantes, en defensa del argumento, detiene en este punto la regresión, no puede quejarse de que Filo, para oponerse al argumento, la dé por terminada con el vegetal. Tal como dice Filo:

Si levanto mi sistema filosófico sobre el primero, prefiriéndolo al segundo, es por mi cuenta y riesgo. Se trata de un asunto, por lo visto, completamente arbitrario. Y cuando Cleantes me pregunte por la causa de esa gran facultad vegetativa o generativa, tendré igual derecho a preguntarle por la causa de su gran principio racional. Ambos hemos convenido en renunciar a estas preguntas, y en este caso, principalmente es en interés suyo no salirse de lo acordado. Juzgando por nuestra limitada e imperfecta experiencia, la generación tiene algunas ventajas sobre la razón, pues todos los días vemos ésta surgir de aquélla; jamás a aquélla de ésta.⁷

Este último argumento resulta bastante inapelable. Cleantes se jacta de la naturaleza «científica» de su razonamiento: un argumento por analogía, fruto de la experiencia. Pero es la misma experiencia la que nos demuestra lo frágil que es la inteligencia y lo mucho que depende de otros factores. De acuerdo con nuestra experiencia, una mente necesita un cerebro, que es una de las apariciones más tardías, inusuales, frágiles y dependientes de la historia de la naturaleza. La «generación», es decir, el nacimiento de una nueva planta o animal a partir de plantas y animales anteriores, resulta, en cambio, de lo más común, y según todo lo que hemos podido observar hasta ahora, es imprescindible para la existencia de la inteligencia. Luego si hay que hablar a partir de la experiencia, resulta altamente improbable que el origen de todo el montaje sea una mente autosuficiente antes que cualquier otra causa física.

Como no parece que haya nada que objetar a las palabras de Filo, será bueno especular un poco acerca del éxito que ha tenido el

argumento del diseño. ¿Por qué ha tenido tan poca aceptación entre la gente la réplica de Filo? Sospecho que el motivo de fondo es el mismo que había detrás de algunos de los problemas que encontramos en el caso de la libertad. Preferimos detener la regresión con la «inteligencia» antes que con la «generación» porque pensamos que nuestra propia experiencia ofrece un ejemplo de un suceso mental incausado, a saber, mi decisión de iniciar una cierta acción, que da lugar a un suceso físico, lo cual nos sirve como modelo para explicar la creación arbitraria del universo por parte de una deidad. Cuando pensamos así olvidamos la observación de Schopenhauer (véase capítulo 3): cuando actuamos, a veces no somos conscientes de los procesos causales, pero no se sigue de ello, y no es cierto, que seamos conscientes de la ausencia de todo proceso causal. Esta relación que se establece entre el argumento del diseño y la concepción intervencionista de la libertad tiene un interesante contenido moral. Se puede decir que las dos imágenes, la de Dios como ser sobrenatural, y la de nuestro «yo» como algo igualmente exterior a la naturaleza, se alimentan mutuamente. Y las dos llevan a la gente a negar el dominio de la naturaleza. Pensamos que «nosotros» dominamos la naturaleza, igual que Dios. Bien al contrario, lo cierto es que nosotros también formamos parte del mundo, una parte muy, muy pequeña.

Ya dije antes que los parecidos vienen casi regalados, y a Filo todavía le quedan unos cuantos por mostrar. Supongamos que nos olvidamos de todas estas objeciones y que le concedemos a Cleantes su «diseñador». ¿Y entonces qué? A veces los diseños son producto de una sola mente. Pero es más frecuente, sobre todo en el caso de diseños de gran complejidad, como por ejemplo los navios, que sean el resultado de la acción conjunta de muchas mentes. Por fin, hay diseños que proceden de mejores diseñadores que otros:

En una palabra, Cleantes, aquel que siga tu hipótesis, puede, quizá,

afirmar o conjeturar que en algún tiempo surgió el universo de algo así como un designio; pero como más allá de esta posición no podrá averiguar ni una sola de las circunstancias, queda, después, en absoluta libertad para fijar todos los puntos de su teología con la mayor licencia de fantasía e hipótesis. Hasta donde él alcanza, es posible que este mundo sea muy defectuoso e imperfecto, si se le compara con un patrón superior, y no es sino el primer y burdo intento de alguna deidad pueril, quien, más tarde, lo abandonó avergonzada de una obra tan imperfecta; o bien es posible que sólo sea la obra de algún subordinado, de alguna deidad inferior, objeto de irrisión para sus superiores; o bien el producto de los muchos años y la chochez de alguna deidad senil, que, desde que ésta murió, ha seguido rodando a toda aventura, a partir del inicial impulso y fuerza activa que ella le comunicó [...] Y yo, por mi parte, no creo que haya motivo alguno para que un sistema teológico tan delirante y poco firme, sea preferible a no tener ninguno.⁸

Lo cual nos lleva inevitablemente a:

El problema del mal

La mayoría de los sistemas religiosos busca algo más en sus dioses que la cualidad tan notablemente abstracta de la «existencia necesaria». Quieren su amor y sus cuidados. Un dios que creara el mundo y luego lo abandonara a sus propios medios no se prestaría mucho a la adoración y tampoco sería una buena fuente de autoridad moral, de modo que los atributos tradicionales de Dios incluyen la perfección moral. Por supuesto, Dios es todopoderoso y omnisciente, pero también es enteramente bondadoso. Y aquí aparece el clásico argumento contra la existencia de Dios: el problema es que en el mundo que él (o ella, o ellos) ha creado, la bondad parece hallarse ausente. Tal como dice Filo:

Admitimos que su poder es infinito: todo lo que dispone se ejecuta; pero ni el hombre ni ningún otro animal es feliz; por lo tanto, no quiere su felicidad. Su sabiduría es infinita: jamás yerra en la elección de los

medios para cualquier fin; pero el curso de la naturaleza no propende hacia la felicidad humana o animal; por lo tanto, no ha sido establecido para ese propósito. En la totalidad del haber del saber humano, no hay inferencias más ciertas e infalibles que éstas. ¿En qué, pues, se parecen su benevolencia y piedad a la benevolencia y piedad de los hombres?

Las viejas preguntas de Epicuro están aún por contestar.

¿Desea evitar el mal, pero es incapaz de hacerlo? Entonces es impotente. ¿Está capacitado, pero no lo desea? Entonces es malévolo. ¿Está deseoso, y tiene capacidad? Entonces, ¿de dónde procede el mal?⁹

El problema de Cleantes es que el mundo tal como lo conocemos es, cuando menos, una media tinta en lo que respecta a la felicidad de sus criaturas. La vida es dura, y para muchos es corta y brutal, llena de deseo y de dolor. El bienestar de muchas criaturas depende de la enfermedad y la muerte de otras. Y es absurdo inferir un creador perfecto a partir de una creación imperfecta. Incluso un padre que sólo fuera medianamente bueno no escogería voluntariamente que su hijo o hija fuera a parar a un entorno brutal si no le costara nada escoger otro mejor. Las propias analogías que proclama Cleantes se vuelven ahora contra él.

Supongamos que nos encontramos en el dormitorio de una escuela o de una universidad. El techo tiene goteras, las ratas se pasean por la habitación, la comida es casi intragable y algunos alumnos llegan incluso a morir de hambre. Detrás de una puerta cerrada se halla el director, pero nunca sale de allí. Nos ponemos a especular acerca de cómo debe de ser el director. ¿Acaso podemos inferir a partir del estado en que hemos encontrado el dormitorio que éste, primero, conoce exactamente las condiciones en que se encuentra; segundo, se preocupa activamente por nuestro bienestar, y tercero, posee una capacidad ilimitada de arreglar las cosas? Tal inferencia sería una locura. Tendríamos prácticamente la certeza de que el director o bien no conoce la situación, o bien no se preocupa por ella, o bien no puede hacer nada para resolverla. Y no mejoraría en nada las cosas el

hecho de que nos encontráramos de vez en cuando con algún alumno que proclamara haber conseguido un acceso privilegiado a la mente del director y que asegurara que éste en realidad sí conoce la situación, se preocupa por ella y tiene los medios y la habilidad necesarias para hacer todo cuanto desee. Nuestra conclusión sería inevitablemente que el alumno se engaña, tal vez a causa precisamente de las privaciones que padece. Las múltiples flaquezas del Windows no han llevado a nadie a la conclusión de que Bill Gates es infinitamente bondadoso, onisciente y capaz de arreglarlo todo.

Se podrían hacer observaciones parecidas a propósito de la creencia de que este mundo es un «valle de lágrimas» o, dicho de otra manera, de que es algo así como una prueba de fuego que sirve para decidir lo que viene después. Los habitantes de nuestro dormitorio bien podrían creer algo parecido: el director estaría observando su comportamiento para distribuirlos al año siguiente en dos dormitorios, uno bueno y otro malo o, mejor dicho, uno perfecto y el otro infernal. Al final podría resultar que estaban en lo cierto. Pero no tienen el más mínimo indicio que les permita creer eso, si se basan en lo que saben hasta ahora. Lo único que pueden saber del director es aquello que ven. Y si él, ella o ello no ofrecen buenas condiciones aquí, ¿por qué habrían de suponer que lo van a hacer en alguna otra parte? Sería como suponer que si allí la temperatura es buena, en algún lugar debe de haber un dormitorio en el que hace mucho calor y otro en el que hace mucho frío. La inferencia no tiene ningún sentido.

Cleantes es especialmente vulnerable ante una objeción de este tipo, ya que él mismo se ha basado en una analogía para inferir de forma razonable la naturaleza del creador a partir de la naturaleza del mundo. Pero incluso si dejamos a un lado las demás objeciones frente al argumento del diseño, lo cierto es que un mundo imperfecto y plagado de taras debería llevarle irremisiblemente a un creador imperfecto y plagado de taras. O bien:

La verdadera conclusión es que la fuente original de todas las cosas es completamente indiferente a todos estos principios, y tan poco le importa el bien sobre el mal como el calor sobre el frío, o la sequía sobre la humedad, o la ligereza sobre la pesantez.¹⁰

Demea —el personaje que se mostró partidario de los argumentos ontológico y cosmológico— se enfrenta a un problema distinto. Él no trataba de inferir su deidad a partir de la naturaleza del mundo, de modo que no es tan vulnerable ante la objeción anterior. La diferencia es que Cleantes deduce la naturaleza de Dios a partir del mundo que todos conocemos, de modo que se ve obligado a demostrar que el mundo es tal como cabría esperar que fuera partiendo de la premisa de un Dios omnisciente, todopoderoso y absolutamente bondadoso. Tiene que demostrar que el mundo es compatible con un ser de estas características. Demea está en condiciones de admitir que el mundo no se ajusta demasiado a nuestras expectativas, y se limita a afirmar que es compatible con su deidad, es decir, que el mundo no refuta la idea de un ser tal.

Sin embargo, Demea todavía tiene que dar respuesta a las «viejas preguntas de Epicuro». Su estrategia, que ha ido ganando en popularidad a lo largo de los siglos posteriores, consiste en refugiarse en la naturaleza misteriosa e incomprensible de la mente divina. Demea se opone a los impíos intentos de comprender la bondad divina a partir del modelo de la bondad humana, o bien las intenciones, percepciones o ideas de Dios a partir de las intenciones, percepciones o ideas de los seres humanos.

En tal caso, el problema consiste en explicar si es posible que tenga alguna consecuencia el hecho de creer en un Dios incomprensible. Tal como diría Wittgenstein más tarde, a propósito de otro asunto:

Una nada presta el mismo servicio que un algo sobre el que nada puede decirse.¹¹

Incluso Hume, el «gran infiel», no tiene el menor problema a la hora de dejar misterios abiertos. Al final de los *Diálogos*, el escéptico *Filo* se permite uno:

Si toda la Teología Natural, como algunos sostienen, se resuelve en una simple, aunque un tanto ambigua, o por lo menos indefinida proposición: que la causa o causas del orden en el universo probablemente guardan alguna analogía remota con la inteligencia humana; si esta proposición no es capaz de extensión, variación, o una más particular explicitación; si no proporciona inferencia alguna que afecte a la vida humana, o que pueda ser motivo para actuar o para abstenerse; y si la analogía, todo lo imperfecta que sea, no es susceptible de llevarse más allá de la inteligencia humana, y no puede transferirse, con alguna apariencia de probabilidad, a las demás cualidades de la mente; si tal es en realidad el caso, ¿qué otra cosa puede hacer el más inquisitivo, contemplativo y religioso de los hombres, que prestar sin reservas su asentimiento filosófico a la proposición, cuantas veces ocurra, y creer que los argumentos en que se funda exceden a las objeciones que pueden ponersele? Seguramente habrá alguna sorpresa, dada la grandiosidad del asunto; alguna melancolía, dada su oscuridad; algún desprecio hacia la razón humana, puesto que no puede proporcionar una solución más satisfactoria, respecto a una tan extraordinaria y magnífica cuestión.¹²

Habiendo ido a parar a una deidad absolutamente misteriosa, Demea se va a encontrar con el problema de no poder sacar de ello ninguna consecuencia. Uno puede dedicarse a explorar la Niebla Misteriosa si lo desea, pero va a salir de ella exactamente igual a como entró. La creencia religiosa, reducida a su núcleo más respetable, se convierte en algo perfectamente inútil. No tiene la menor consecuencia.

Esto es algo que sorprende a la gente, hasta el punto de que ha llevado a muchos comentaristas a discutir largamente acerca de si Hume era en realidad un teísta o un ateísta. Muchas personas creen

que la diferencia entre ser un teísta, alguien que cree en Dios, o un ateuista, alguien que no cree en Dios, tiene una importancia extraordinaria. Pero se desvanece cuando pensamos que una nada presta el mismo servicio que un algo sobre el que nada puede decirse. Si todo cuanto podemos creer de forma razonable es que la causa del universo mantiene algu0n tipo de analogía remota e inconcebible con las dem0s operaciones que se dan en la naturaleza, entonces nos quedamos con una idea perfectamente in0til, y volvemos de aquellas misteriosas regiones sin haber logrado ninguna comprensi3n real de las cosas. Podr0amos hacernos eco de la observaci3n de Wittgenstein y decir que Hume ha «desconstruido» la aparente diferencia entre el te0simo y el ate0simo.

En particular, si la «bondad divina» no debe ser entendida en t0rminos de aquello que nosotros consideramos bueno (de modo que, por ejemplo, ser0a «bueno» en este otro sentido que Dios enviara la plaga bub3nica a unos ni0os indefensos), no afecta al modo en que deber0a vivir mi vida. No me da ninguna pista m0s acerca de si deber0a preferir el placer o el dolor, o si deber0a poner la otra mejilla o tomar el ojo por el ojo, de las que tengo para preferir el calor al fr0o. Y sin embargo se supone que eso es lo que hace la religi3n. Su importancia reside en que la gente acostumbra a pensar que es relevante a la hora de definir nuestro modo de actuar. Pero ahora descubrimos que si llevamos los argumentos tradicionales hasta sus 0ltimas consecuencias, resulta que es perfectamente irrelevante a estos efectos.

La teodicea es una rama de la teolog0a que trata de encontrar una respuesta al problema del mal. Una posibilidad es decir que hay bienes que parecen presuponer ciertos males. Podemos tratar de dar 0nimos a la gente del dormitorio sucio y desastrado exaltando las virtudes de la paciencia y la entereza, unos bienes que florecen tan s3lo en medio de la miseria y la adversidad. El problema que tiene este argumento es que a pesar de todo seguimos pensando que las

cosas mejoran cuando las situaciones que fomentan tales virtudes pierden algo de su crudeza. Las imperfecciones de Windows han contribuido indudablemente a la extensión de las virtudes de la paciencia o la entereza, pero ni siquiera Microsoft ha empleado este argumento para defender su producto, y ése es el motivo de que sigan intentando mejorarlo.

A veces la gente trata de resolver el problema volviendo otra vez a la creencia en una deidad genuinamente buena, y buena en un sentido que podemos comprender, mediante lo que se conoce como el «argumento de la libertad». La idea es que Dios creó un universo bueno, y que por su bondad nos creó a nosotros libres. Pero a causa del uso impropio que hemos hecho de la libertad que nos fue otorgada, nosotros mismos hemos traído el mal a un mundo que de otro modo hubiera seguido siendo bueno. La idea queda resumida en el mito de la Caída y la expulsión del Paraíso.

Se pueden plantear muchas objeciones a este argumento. En primer lugar, parece depender de una concepción de la libertad que resulta incoherente: la concepción intervencionista según la cual hay algo que no forma parte del orden natural (el Verdadero Yo) que en ocasiones interfiere sobre él. De otro modo, si nuestra libertad fuera concebida en términos compatibilistas, nuestras decisiones tendrían un carácter natural y por lo tanto se remontarían, en último término, a Dios. Si éste no hubiera querido que Stalin masacrara a millones de personas, no debería haber creado esta naturaleza que ha dado lugar finalmente a los módulos de decisión de un individuo así.

En segundo lugar, no es en absoluto cierto que la totalidad de los males que afligen a la humanidad, y ni siquiera una parte significativa de ellos, guarden la más mínima relación con las decisiones de las personas. Son el resultado de la enfermedad, del dolor, del deseo o bien de un mero accidente. Afligen tanto a los animales como a los seres humanos, y lo hicieron desde mucho antes de la aparición de estos últimos.

En tercer lugar, incluso si aceptáramos la metafísica de la libertad, cabría esperar de un buen Dios que protegiera al menos a algunos de los más débiles de los abusos de la libertad de los más fuertes. Un padre podría reconocer el valor de permitir que los niños tomen sus propias decisiones y darles por ello cierta libertad. Pero si alguno de los mayores demostrara una alarmante tendencia al asesinato y a la mutilación de los más jóvenes, sería razonable por parte del padre que lo pusiera bajo vigilancia, o bien que protegiera al menor impidiendo que el otro llevara a cabo sus planes. Por desgracia, Dios no actúa de este modo en el mundo que conocemos. No existen guarderías naturales donde los más débiles queden separados de los más fuertes. Somos nosotros quienes debemos esforzarnos por crear nuestros espacios seguros.

Mi opinión personal sobre esta cuestión es que las tradiciones religiosas muestran su mejor rostro cuando abandonan las virtudes clásicas de Dios. En algunas tradiciones la divinidad se sitúa por encima de la bondad y la virtud o, para emplear la expresión más mundana de Hume, no tiene mayor preferencia por el bien en lugar del mal de la que tiene por el calor en lugar del frío. En otras tradiciones, no se le considera en absoluto omnipotente, sino que se halla sujeto a fuerzas que no han sido creadas por él. Cualquiera de ellas permite al menos algún tipo de teodicea. Pero si fuera realmente importante para nosotros desentrañar la naturaleza de la mente divina a partir de la naturaleza de su creación, deberíamos considerar seriamente la idea de que él (ella, ellos o ello) tiene un sentido del humor bastante retorcido. Después de todo, tal como dice un chiste judío, permitió que su pueblo escogido diera vueltas por el desierto durante cuarenta años sólo para que fueran a parar al único punto del Próximo Oriente donde no hay petróleo.

Los milagros y su testimonio

Tal vez sea cierto que los argumentos de fondo que hemos examinado hasta ahora fracasan. Sin embargo, muchas personas siguen pensando que la fe religiosa encuentra un buen fundamento en la ocurrencia de sucesos milagrosos. Un profeta nos da garantías de la existencia de Dios cuando es capaz de predecir el futuro, aparecer después de muerto, realizar curaciones milagrosas y otras cosas por el estilo.

La mayoría de nosotros no ha tenido el privilegio de observar en directo tales sucesos. Nuestra creencia procede más bien de otras fuentes: del testimonio. Hemos tenido noticia de ellos a través de la Biblia, el Corán, las *Vidas de los Santos* o, incluso, el *National Enquirer*. No hemos visto nunca personalmente, por ejemplo, que un miembro amputado vuelva a crecer hasta recuperar su tamaño normal, pero tal vez hemos oído hablar de un caso parecido cuya certeza estaría fuera de toda duda y que habría ocurrido en algún lugar más allá de las montañas. La mayoría de las personas no ha sido abducida por los marcianos, pero puede ser que crea de todo corazón en el relato de aquellos que aseguran haber vivido la experiencia, o sitio sus hermanos o sus primos. Incluso si no hemos visto recientemente a Elvis, enterrado hace ya bastantes años, podríamos leer que otras personas lo han visto y creerlo.

Fue Hume quien planteó la pregunta fundamental: ¿cuándo resulta razonable creer en tales testimonios?

Supongamos que dejamos a un lado el aspecto «milagroso», es decir, la cuestión de si alguno de estos sucesos se debe a la acción de poderes invisibles o a la intervención divina. Sigue siendo cierto que cualquier hecho que aspire a ser considerado un milagro debe resultar sorprendente en grado superlativo, el tipo de cosa que en el curso normal de los acontecimientos simplemente no llega a ocurrir nunca (no nos interesa aquí la cuestión de si la creación entera podría considerarse milagrosa, puesto que eso nos llevaría de vuelta al argumento cosmológico). Para dar fe de una intervención divina, no

basta con tener una participación destacada en sucesos inusuales. Se requieren hechos realmente increíbles: personas que se elevan por el aire, plomo que flota, agua que se convierte en vino, muertos que vuelven a la vida. El reto al que se enfrenta el potencial hacedor de milagros vendría a ser: «Adelante, impresioname». Así pues, ¿cuándo resulta razonable creer en el testimonio de tales sucesos insólitos, completamente ajenos a la normalidad?

Hume comienza por establecer un principio bastante razonable acerca de lo que dicen las personas. Es un hecho, pensamos, que la mayoría de las veces dicen la verdad. Hume afirma que cuando realizamos una inferencia desde una premisa del tipo «Tal persona me dice *que p*» hasta la conclusión «Luego *p* es probablemente verdadero», estamos haciendo algo muy parecido a lo que hacemos cuando inferimos a partir de un hecho como «La bola de béisbol se dirige a la ventana» otro hecho como «La ventana se va a romper». Tales inferencias son empíricas (*a posteriori*), y se basan en el tipo de relaciones que nuestra experiencia encuentra en el mundo. La veracidad de los testimonios de la mayoría de las personas es un hecho que viene respaldado por nuestra experiencia. Pero cuando las cosas se ponen difíciles no confiamos tanto en ellos. Puede ser que la evidencia sea contradictoria, es decir, que algunos hechos apunten en una dirección y otros en otra:

En el presente caso esta oposición entre evidencias puede derivar de diversas causas: de la oposición del testimonio contrario, del carácter y número de los testigos, de la manera de dar su testimonio o del conjunto de todas estas circunstancias. Dudamos de una cuestión de hecho cuando los testigos se contradicen, cuando son sólo pocos o de carácter dudoso, cuando tienen intereses en lo que mantienen, cuando atestiguan con vacilaciones o, por el contrario, con aseveraciones demasiado violentas. Hay otros muchos detalles de esta clase que pueden disminuir o destruir la fuerza de cualquier argumento que se deriva del testimonio humano.¹³

En otras palabras, la propia experiencia nos enseña en qué momentos debemos evitar ser demasiado crédulos. Pero ahora supongamos que el testimonio ofrecido es completamente insólito, próximo a lo milagroso. En tal caso:

El mismo principio de la experiencia, que nos da cierta seguridad en el testimonio de los testigos, en este caso también nos da otro grado de seguridad en contra del hecho que ellos intentan establecer. De esta contradicción necesariamente surge un contrapeso y la consiguiente destrucción mutua de creencia y autoridad.¹⁴

Antes de detenernos a examinar este razonamiento, deberíamos ver a dónde nos lleva. Hume extrae una famosa conclusión:

La simple consecuencia es (y trátase de una máxima general digna de nuestra atención) «que ningún testimonio es suficiente para establecer un milagro, a no ser que el testimonio sea tal que su falsedad fuera más milagrosa que el hecho que intenta establecer; e incluso en este caso hay una destrucción mutua de argumentos, y el superior sólo nos da una seguridad adecuada al grado de fuerza que queda después de deducir el inferior». Cuando alguien me dice que vio resucitar a un muerto, inmediatamente me pregunto si es más probable que esta persona engañe o sea engañada, o que el hecho que narra haya podido ocurrir realmente. Sopeso un milagro en contra de otro y, de acuerdo con la superioridad que encuentro, tomo mi decisión y siempre rechazo el milagro mayor. Si la falsedad de su testimonio fuera más milagrosa que el acontecimiento que relata, entonces, y no antes, puede pretender obtener para sí mi creencia y opinión.¹⁵

El argumento se puede analizar de diversas maneras. Puede ser útil pensarlo del siguiente modo:

Supongamos que alguien me habla de un suceso (*m*) altamente sorprendente o improbable. Sea *m* el suceso más improbable que seamos capaces de imaginar. Así que la evidencia en favor de *m* es que «esta persona dice que *m* sucedió». En este punto puedo escoger

entre dos posibles planteamientos acerca de la cuestión:

- a)Esta persona dice que *m* sucedió. Pero *m* no sucedió.
- b)Esta persona dice que *m* sucedió. Y *m* sucedió.

Ahora bien, tanto (a) como (b) contienen un elemento sorprendente. El planteamiento (a) contiene la siguiente sorpresa: esta persona mintió. El planteamiento (b) contiene la sorpresa de que *m* sucediera. De modo que debo sopesar cuál de los dos es más sorprendente o improbable, y luego rechazar «el milagro mayor».

El problema, tal como anuncia alegremente Hume, es que los testimonios falsos son bastante comunes. El caso más evidente es el de una mentira deliberada. También hay casos de alucinaciones. A veces se producen espectaculares lapsus de memoria. Siempre que haya intercambios de información, se van a producir errores: malentendidos, traducciones desviadas, casos en los que se toma una expresión empleada en sentido metafórico como si fuera una verdad en sentido literal, y así sucesivamente. Luego (a) no es tan improbable como (b). El planteamiento (b) implica la presencia de un milagro: el suceso más improbable que pueda imaginarse. El planteamiento (a) implica algo que sabemos que se produce constantemente: la gente se equivoca al interpretar las cosas. El principio de que «ningún testimonio es suficiente para establecer un milagro, a no ser que el testimonio sea tal que su falsedad fuera más milagrosa que el hecho que intenta establecer» constituye un listón increíblemente difícil de superar para cualquier testimonio. E incluso si alguno lo consiguiera, nos dejaría sumidos en una especie de confusión: no sabríamos qué creer, de modo que la opción más inteligente sería suspender el juicio.

En realidad, prosigue Hume, ninguna de las evidencias que se han empleado para fundar un sistema religioso ha logrado siquiera acercarse al listón. Esta afirmación se apoya en una serie de

observaciones: los relatos de los milagros acostumbran a proceder de épocas o regiones remotas o incivilizadas; de personas agitadas por las pasiones; o de personas que tienen un interés positivo en vender su historia:

Los sabios prestan una fe muy académica a todo informe que favorece la pasión del narrador, bien que ensalce su país o su familia o a sí mismo o que, de cualquier otro modo, coincida con sus inclinaciones y propensiones naturales. Pero ¿qué mayor tentación que presentarse como un misionero, profeta o embajador del cielo? ¿Quién no se enfrentaría con muchos peligros y dificultades con tal de alcanzar un carácter tan sublime? O si, con la ayuda de la vanidad o una imaginación calenturienta, un hombre, primero, se convirtiese y cayera, seriamente en el engaño, ¿tendría después escrúpulos en utilizar fraudes píos en apoyo de una causa tan meritoria y sagrada?¹⁶

Observa también que la gente está muy aficionada a estos relatos:

La afección de sorpresa y de asombro que producen los milagros, al ser una emoción agradable, provoca una fuerte propensión a creer en los acontecimientos de los que se deriva. Y esto va tan lejos que, incluso los que no pueden disfrutar este placer inmediatamente, ni pueden creer en los acontecimientos milagrosos de los que se les informa, sin embargo, gustan de participar en la satisfacción de segunda mano o, de rechazo, encuentran placer y orgullo en excitar la admiración de otros.

Con avidez se acogen los relatos milagrosos de viajeros, sus descripciones de monstruos terrestres y marinos, sus narraciones de aventuras maravillosas, de hombres extraños y de rudas costumbres. Pero si el espíritu religioso se une al gusto por el asombro, se acabó el sentido común, y en esta situación pierde el testimonio humano todas sus pretensiones de autoridad.¹⁷

Y todavía introduce un comentario más sutil acerca de la relación entre las diversas religiones, cada una de las cuales tiene su propio surtido de milagros:

[C]onsideremos que, en cuestiones de religión, lo que es distinto también es contradictorio y que es imposible que las religiones de la antigua Roma, de Turquía, de Siam y de China pudieran, todas ellas, establecerse sobre un fundamento sólido; por tanto, todo milagro que se pretende que ha tenido lugar en cualquiera de estas religiones (y todas ellas abundan en milagros), como su consecuencia inmediata es establecer el sistema particular al que se atribuye, por tanto, tiene la misma fuerza, al menos directamente, para desautorizar a los demás sistemas. Al destruir un sistema rival, igualmente destruye el crédito de los milagros sobre los que este sistema se estableció, de modo que los prodigios de las distintas religiones han de considerarse como hechos contrarios, y las evidencias de estos prodigios, fuertes o débiles, como opuestas entre sí.¹⁸

También sería ésta la respuesta de Hume si le objetaran que es imposible que tanta gente esté equivocada. Se mire por donde se mire, sigue habiendo mucha gente que necesariamente se equivoca.

El argumento que emplea Hume en esta ocasión es maravillosamente económico. Un filósofo dotado de menos sutileza habría intentado llegar a algún tipo de conclusión metafísica, como por ejemplo la absoluta imposibilidad de los milagros. Hume no necesita esta conclusión, y ni siquiera la persigue. Desde el punto de vista metafísico, concede la posibilidad de una intervención divina. Podría ser que hubiera una divinidad que en alguna ocasión podría haber permitido que alguien caminara sobre las aguas o alimentara a cinco mil personas con unos cuantos panes y peces. Sin embargo, la experiencia sigue siendo nuestra única guía para saber si estos hechos ocurren realmente. Si hemos de creer que sí a causa de un testimonio, tiene que tratarse de un buen testimonio: muy bueno, milagrosamente bueno en realidad. Pero nunca conseguimos encontrar los testimonios adecuados.

La gente que escucha por primera vez el argumento de Hume tiende a pensar que es indebidamente cínico, ya que revelaría una

actitud de desconfianza y suspicacia frente a los relatos de otras personas. Personalmente no pienso que sea así, o por lo menos no creo que la sospecha vaya más allá de lo que avalan las tendencias manifiestas de las personas. Después de todo, hay que ser extremadamente inocente para negar, por ejemplo, que es bueno sospechar de aquellos relatos que satisfacen la vanidad de quien los relata. Veamos la siguiente cita extraída del periódico británico *The Independent*, en relación con un informe presentado por el Real Colegio de Psiquiatras:

Según el Real Colegio de Psiquiatras, uno de cada seis de nosotros es un neurótico. También deben pensar que todos nosotros somos unos ingenuos. Presenta un informe: la forma políticamente correcta de anunciar tus servicios. ¿Y luego qué? ¿El Instituto de Constructores dirá que siete de cada diez casas deben ser reconstruidas, o la Asociación de los Mecánicos que trece de cada veinte coches deben ser reparados?

En realidad, los temas que trata Hume en la segunda parte de su gran ensayo anticipan todo un campo de estudios académicos. Los psicólogos investigan actualmente las distorsiones cognitivas más comunes: los errores en la percepción o en la memoria, la influencia de otras personas, las insidiosas virtudes de la autoconfianza o el gusto por lo maravilloso se consideran factores que interfieren en la capacidad de las personas para distinguir lo verdadero de lo falso. Somos unos instrumentos bastante buenos para registrar la verdad y rechazar la falsedad, pero no lo somos tanto como nos gusta creer, y a menudo no lo somos en absoluto.

El argumento de Hume se puede formular elegantemente en los términos del teorema de Bayes, que será expuesto en el próximo capítulo. Tal vez el lector prefiera volver sobre este punto una vez que haya asimilado la explicación que se ofrece allí. En términos de Bayes, sea h la hipótesis de que ha ocurrido un milagro, y e el hecho de que una persona o un grupo de personas dicen que ha ocurrido. La

probabilidad previa o antecedente de que el milagro haya ocurrido es muy, muy baja. Su «probabilidad de base» se acerca a cero. Ello es así porque los milagros son la clase de cosas que no ocurren nunca, o casi nunca. Cada mañana, cuando salgo para ir a la oficina mi esposa me previene contra el frío, o el tráfico o mis compañeros de trabajo. En cambio, no me previene contra los elefantes voladores, la posibilidad de ser tomado como esclavo sexual por los marcianos o una posible conversación con un Elvis redivivo. Consideremos ahora el hecho de que existe alguien o algún texto que afirma que el milagro efectivamente se produjo. Pues bien, lamentablemente esto último sí que es la clase de cosa que acostumbra a ocurrir. La probabilidad previa o antecedente de que se presente una evidencia tal no ha sido nunca demasiado baja, ya que existe un buen número de hipótesis distintas y perfectamente naturales capaces de explicarlo. Se trata de las clásicas debilidades humanas: el engaño, la ilusión, el desorden de las pasiones, el error, y así sucesivamente. Incluso los defensores de una determinada colección de milagros tienen que creer en tales debilidades para poder deshacerse de posibles impostores. La Iglesia romana tiene un departamento entero dedicado a la tarea de desenmascarar falsos milagros. Para los cristianos es mejor no creer que Mahoma tomara realmente un vuelo nocturno desde La Meca hasta Jerusalén, puesto que las pretensiones de éste como hacedor de milagros son antagónicas a las de Jesús. Pero eso significa que la probabilidad previa o antecedente de e es relativamente alta. Hay muchas formas de generar «pruebas falsas». Tal como veremos, Bayes exige que consideremos las probabilidades previas o antecedentes antes de determinar la probabilidad de una hipótesis en función de la evidencia presentada. Lo ideal sería una hipótesis que no resultara para nada improbable y una evidencia que difícilmente se pudiera producir a menos que la hipótesis fuera cierta. Pero en este tipo de casos la probabilidad previa o antecedente se halla exactamente en la situación contraria. La hipótesis es

enormemente improbable y es fácil que la evidencia venga motivada por otras circunstancias, de modo que el cálculo de Bayes siempre se pronunciará en contra de la verdad del testimonio y en favor de la uniformidad de la naturaleza.

No quiero decir con todo esto que cualquier relación de circunstancias en principio extrañas a nuestra experiencia deba ser necesariamente falsa. La ciencia avanza precisamente gracias a este tipo de cosas. Pero razonamos de la forma adecuada cuando mantenemos una actitud escéptica hasta que el nuevo fenómeno se repita y quede establecido como una más de las regularidades de la naturaleza.

Cuando pasamos a la teología de los milagros, la situación se vuelve aún peor. Una divinidad que impusiera un conjunto de leyes a la naturaleza y las mantuviera de forma inflexible poseería al menos cierta dignidad. Otra que permitiera saltos e interferencias y que realizara algunos trucos para ser glorificada resultaría menos atractiva. ¿Por qué motivo tendría que haber milagros? No son precisamente lo que cabría esperar. Un par de milagros que ventilaran a todos los Hitlers y Stalins resultarían bastante más útiles que uno destinado a convertir el agua en vino con ocasión de un particular festejo de bodas. Es sin duda un gesto apreciable por parte de Dios hacer que san José levite ante un retrato suyo, pero uno habría preferido, por decir algo, la milagrosa puesta en cuarentena o destrucción del virus del sida. Eso es al menos lo que en un principio cabría esperar, teniendo en cuenta que el mundo está regido por un buen Dios. Pero el mundo que conocemos no confirma nada de todo esto. Pronto veremos que este razonamiento también se puede analizar en los términos de Bayes. En este caso, el punto flaco es el grado de correspondencia que pueda haber entre la evidencia y la hipótesis, la segunda de las tres figuras fundamentales del teorema de Bayes.

Infinito nada

Ninguno de los argumentos metafísicos que hemos considerado hasta ahora parece capaz de confirmar la hipótesis de que el universo sea la creación de un Dios en el sentido tradicional. Por otro lado, el análisis de Hume acerca de los testimonios de los milagros concluye que no tienen el menor valor como evidencia. La fe religiosa deberá buscar nuevos argumentos si quiere dar respuesta a todas estas carencias.

El matemático y teólogo francés Blaise Pascal (1632-1662) supo encontrar una respuesta interesante e ingeniosa, conocida como la apuesta de Pascal.¹⁹ A diferencia de los argumentos que hemos visto hasta ahora, no pretende ser un argumento en favor de la verdad de la fe religiosa, sino de la utilidad que puede tener la creencia en alguna de las versiones del monoteísmo, sea la judía, la cristiana o la islámica.

El argumento es el siguiente. En primer lugar, Pascal confiesa su ignorancia metafísica:

Hablemos ahora según las luces naturales.

Si existe un Dios, es infinitamente incomprensible, puesto que, no teniendo ni partes ni límites, no tiene relación alguna con nosotros. Somos, por consiguiente, incapaces de conocer ni lo que es, ni si es. Siendo esto así, ¿quién se atreverá a intentar resolver esta cuestión? No seremos nosotros, que no tenemos ninguna relación con él.

No resulta demasiado claro por qué la excusa se refiere sólo a los cristianos, y excluye a las demás religiones, así como a la creencia en las hadas, los fantasmas, Elvis redivivo y L. Ron Hubbard. No obstante, supongamos que la elección se plantea entre la fe religiosa y una vida de duda y escepticismo acerca de la religión:

[E]s preciso apostar. No es voluntario, estáis embarcados. ¿Por cuál os decidiréis, por tanto? [...] Pesemos la ganancia y la pérdida apostando cruz a que Dios existe. Tengamos en cuenta estos dos casos: si ganáis,

ganáis todo, y si perdéis, no perdéis nada: apostad, pues, a que El existe, sin vacilar.

Pascal es lo bastante lúcido como para darse cuenta de que es un motivo más bien raro para escoger una creencia. Pero luego añade, con gran perspicacia, una observación:

[V]uestra incapacidad para creer procede de vuestras pasiones[,] puesto que la razón os lleva a ello, y sin embargo, no podéis conseguirlo. [...] [Aprended] de los que han estado maniatados como vosotros y que apuestan ahora todo su bien. [...] [S]eguid el modo con el que ellos han comenzado. Consiste en hacer todo como si creyesen, tomando agua bendita, encargando misas, etc. Naturalmente, eso mismo os hará creer y os embrutecerá.

Después de un «lavado de cerebro» de este estilo, uno se convierte en un creyente. Y es entonces cuando se cosechan los frutos de la fe: sus infinitos frutos, en caso de que exista nuestro Dios. ¿Y si no fuera así? Pues bien, hemos perdido muy poco, si lo comparamos con la eternidad: tan sólo lo que Pascal llama los «placeres malsanos», como podría ser ir a jugar al golf los domingos, en lugar de ir a misa.

Es habitual presentar el argumento en forma de una tabla de dos por dos con las siguientes opciones:

	Dios existe	Dios no existe
Creo en él	i+ infinito!	0
No creo en él	i-infinito!	0

Los ceros de la parte derecha corresponden a la idea de que no se pierde mucho en esta vida por el mero hecho de creer o no creer. La vida presente tiene un valor despreciable en comparación con las promesas que reciben los creyentes. El infinito con el signo « + » corresponde a la gloria infinita. El infinito con el signo «—» que hay en la parte inferior izquierda corresponde al celoso Dios de la tradición,

que manda al infierno a aquellos que no creen en él y que por supuesto anima a sus seguidores a que les hagan pasar un mal rato también en este mundo. Pero también podríamos ignorar la opción del «-infinito». Incluso si pusiéramos un O en su lugar, la apuesta sería buena. Lo seguiría siendo incluso si Dios no castigara la falta de fe, ya que el fantástico premio del «+infinito» deja la elección prácticamente resuelta. En términos de la teoría de la decisión, la opción de creer en Dios «domina», puesto que puede ganar y, en cambio, no puede perder. Así pues: ¡a por ella!

Por desgracia, el error que comete este argumento parece muy simple, una vez que se ha descubierto.

Pascal parte de una postura de ignorancia metafísica. No sabemos nada acerca de lo que pueda haber más allá de la experiencia. Pero el planteamiento de la apuesta supone que sí sabemos algo. Pretende conocer los premios y los castigos que van asociados a la fe en un Dios cristiano. Se trata de un Dios que agradece y premia el hecho de que vayamos a misa y que se mostrará indiferente o, de acuerdo con la opción del «-infinito», seriamente disgustado por nuestra falta de asistencia. Pero las opciones son falsas, ya que si somos realmente ignorantes desde el punto de vista metafísico, debemos aceptar que es al menos igual de probable que las opciones queden distribuidas del siguiente modo:

En verdad existe un Dios inmensamente poderoso y benevolente, el cual (o la cual, los cuales o lo cual) ha dejado establecido lo siguiente: los seres humanos buenos son aquellos que se dejan guiar por la luz natural de la razón, que les ha sido entregada para que puedan controlar sus creencias. Estos seres humanos buenos hacen caso a los argumentos, y evitan por lo tanto las creencias religiosas. Aquellos que demuestren la suficiente fortaleza de ánimo como para no creer en tales cosas irán al cielo. Los demás irán al infierno.

Esta divinidad no nos resulta sin duda tan familiar como el celoso

Dios de la tradición que se preocupa sobre todo por si la gente cree en él. (¿Por qué habría de ser tan celoso Dios? ¿Acaso no podría tratarse de una simple proyección de las ambiciones y las pasiones sectarias de la humanidad? ¡Estás con nosotros o contra nosotros! El escéptico francés Voltaire decía que Dios creó al hombre a su imagen, y la humanidad le devolvió el favor.) Pero el problema de Pascal consiste en que si es verdad que no sabemos nada, entonces no sabemos si el escenario que acabamos de describir es menos probable que la versión cristiana planteada por él. Apuesto a que este Dios que castiga la fe en lugar de la falta de fe es igual de probable que el anterior y, por cierto, mucho más razonable.

Y tal como hizo Hume, también podemos criticar el hecho de que Pascal reduzca la cuestión a dos opciones: ir a misa o no creer. En un contexto más amplio habría que añadir muchas más: el Corán o la misa, L. Ron Hubbard o Swami Maharishi, y la *Comunidad de los conceptos de Aquario y del gobierno del nuevo orden divino* o la *Primera iglesia de Internet*. La apuesta tiene que guardar silencio en relación con estas otras opciones.

El deseo de creer

Podemos pasar ahora a considerar brevemente la propuesta «fideísta», que reconoce la falta de base de los argumentos, aunque defiende el derecho de las personas a creer en lo que quieran. El fideísmo atribuye cierto mérito a la fe ciega, parecido al de la madre que se niega a admitir la culpabilidad de su hijo a pesar de las pruebas que le condenan.

Un filósofo entregado profesionalmente a la verdad y a la razón no puede ser demasiado favorable a esta actitud. Una fe que desafíe a la razón puede ser vista como una bendición por aquellos que la comparten, pero fuera de su círculo será tachada de credulidad y superstición, y se verá en ella la amenaza del fanatismo y la

intolerancia. En el segundo capítulo del famoso ensayo *Sobre la libertad*, John Stuart Mill (1806-1873) evoca de forma memorable la atmósfera de «esclavitud mental» que se crea allí donde falta el cuestionamiento de la razón crítica. Incluso la verdad, dice Mill, si es mantenida como un prejuicio ajeno e impermeable al razonamiento, «no es más que otra superstición, que accidentalmente se expresa con las mismas palabras que la verdad» .²⁰ En un texto clásico del escritor inglés de finales del siglo XIX W. K. Clifford se comparan las creencias basadas en una evidencia insuficiente con los placeres robados.²¹ Podemos encontrar una cita adecuada en Samuel Taylor Coleridge:

Aquel que empieza por amar la Cristiandad antes que la verdad, pronto amará su propia secta o Iglesia antes que la Cristiandad, y terminará por amarse a sí mismo antes que todo lo demás.²²

Sin embargo, por muy atractivas que sean estas ideas, resulta difícil demostrar que el hábito de conservar una fe ciega sea necesariamente tan pernicioso. Si después de llegar a la conclusión de Hume de la inutilidad de la fe, la queremos llenar de esperanzas, miedos, propósitos, y con todos los colores de nuestro credo particular, ¿qué hay de malo en eso? ¿Acaso no es una buena cosa la simple devoción?

Ciertamente, hay personas que piensan que las creencias arbitrarias son una buena cosa. Ahora mismo tengo delante el anuncio de una empresa que se define a sí misma como «su supermercado metafísico». Está especializada en libros y música de la New Age, esencias florales y aceites esenciales, aromaterapia, terapia magnética, terapia de equilibrio lumínico, astrología y numerología, tarot y lectura de cartas, cristales y piedras preciosas y, por último, como una lastimosa nota que recuerda el sano juicio, hierbas medicinales. ¿Qué motivo pueden tener los pensadores para ridiculizar las sencillas devociones de la gente?

Por supuesto, hay devociones sencillas que no entran dentro de esta norma genérica de respeto. Si después de internarme en la Niebla Misteriosa regreso convencido de que Dios me ha ordenado matar a todas las chicas jóvenes, o a las personas que tienen la piel del color equivocado, o a la gente que asiste a la iglesia equivocada, o a la gente que practica el sexo de la forma equivocada, eso ya no está tan bien; de modo que tendremos que aplicar nuestros valores humanos, nuestro propio sentido de lo bueno y de lo malo, de lo justo y de lo injusto, para distinguir si regreso de la montaña convertido en una persona admirable o en un lunático.

Parece que hemos ido a parar inevitablemente al terreno de la ética. Unas pocas líneas no son suficientes para hacer una lista completa de las virtudes y los males de las creencias religiosas, del mismo modo que resulta difícil (aunque no imposible) valorar el beneficio o el daño que produce la creencia en la terapia magnética, el Feng Shui o lo que sea. Es indudable que cumplen cierta función: dar respuesta a una serie de deseos y necesidades humanos. Algunas de estas necesidades forman parte del bagaje común de la humanidad: ya he hablado de la necesidad que tenemos de ciertas ceremonias en algunos momentos cruciales de la vida, o de la necesidad de poesía, símbolos, mitos y música para expresar emociones y vínculos sociales que sentimos la necesidad de expresar. Todo esto es bueno. Por desgracia, algunos de los deseos resultan menos admirables: el deseo de levantar barreras que nos separen de los demás, de imponerles nuestra forma de entender la vida, de buscar justificaciones morales para el colonialismo, o para el imperialismo tribal o cultural; todo ello queda libre de culpa si se hace en el nombre del Señor. Por cada místico de talante pacífico y caritativo hay un capellán castrense que se esfuerza por convencer a las tropas de que Dios está de su parte. Personalmente no he visto nunca una pegatina en un coche que dijera: «Si amas a Jesús, odia», pero a veces me pregunto por qué no. Sería un buen eslogan para la

moral religiosa.

Tal vez sorprenda que la cuestión se haya convertido en un problema práctico o moral. Podría parecer que se trata de un simple enfrentamiento entre la razón (buena) y la fe (mala o, al menos, sospechosa). Pero el propio Hume se encargó de pintar un cuadro más complicado. Por razones que pronto tendremos ocasión de ver, parece que el sentido común que aplicamos cotidianamente se halla plagado de elementos de mera confianza o fe. Ya en el capítulo 1 descubrimos nuestra fe en el mundo exterior o en el pasado. En los próximos dos capítulos veremos otros puntos en los que, tal como Hume fue el primero en señalar, la seguridad con la que nos movemos en el día a día parece estar basada en la fe antes que en la razón.

Evidentemente, el hecho de que adoptemos una actitud hacia el «fideísmo» que permita que ciertas creencias religiosas concretas escapen a la razón depende en gran medida de los resultados que recientemente hayamos obtenido de ello. Hume nació cuando todavía no habían pasado veinte años desde la última ejecución oficial por motivos religiosos que tuvo lugar en Gran Bretaña, y él mismo se vio acosado por la entusiasta hostilidad de los creyentes. Si en la época y en el lugar donde nos toca vivir lo único que vemos son excursiones parroquiales y obras benéficas, no tendremos motivos para preocuparnos demasiado. Pero hay mucha gente que baja de la montaña pertrechada con sus propias certezas prácticas acerca de cómo deberíamos ser.

Tal vez algún día descubriremos el modo de satisfacer nuestras necesidades sin alimentar nuestros bajos instintos, pero la historia de la humanidad sugiere que no es muy prudente confiar en ello.

CAPÍTULO

6

Razonar

El presente capítulo pretende que nos familiaricemos con algunas de las categorías básicas que se emplean para analizar nuestra forma de razonar. Queremos que nuestros razonamientos sean buenos. Queremos encontrar métodos fiables para distinguir lo verdadero de lo falso y configurar nuestras creencias acerca del mundo. Pero, ¿cuáles son estos métodos fiables y cómo podemos reconocerlos como tales? En este capítulo daremos un breve repaso a la lógica formal, y estudiaremos algunos de los problemas que presenta el razonamiento inductivo, así como algunos aspectos del razonamiento científico.

Algo de lógica

Los primeros elementos operativos de un argumento son *las premisas*. Constituyen el punto de partida, o aquello que se da por aceptado o asumido en relación con el argumento. Un argumento puede tener una o varias premisas. A partir de las premisas del argumento se deriva una *conclusión*. Cuando reflexionamos acerca de un argumento, tal vez porque nos resistimos a aceptar la conclusión,

tenemos dos opciones. En primer lugar, podemos rechazar una o varias premisas. Pero, en segundo lugar, también podemos rechazar la forma en que se ha derivado la conclusión a partir de las premisas. La primera reacción supone que alguna de las premisas es falsa. La segunda, que el argumento es inválido. Por supuesto, un argumento puede estar sujeto a ambas críticas: que sus premisas sean falsas y que el razonamiento a partir de ellas sea inválido. Pero se trata de dos críticas distintas (y las dos palabras, «falso» e «inválido», la subrayan adecuadamente).

En la vida cotidiana, los argumentos reciben también otra clase de críticas, basadas en otro tipo de consideraciones. Puede que las premisas no sean demasiado razonables. Es absurdo construir un elaborado argumento a partir de la premisa de que la próxima semana voy a ganar la lotería, cuando es algo tan probable como que se me caiga un piano en la cabeza. A menudo resulta inadecuado dar por válidas unas premisas que son ellas mismas objeto de controversia. En determinadas circunstancias puede ser grosero y ofensivo argumentar ciertas cosas. Sin embargo, «lógico» no es sinónimo de «razonable». La lógica se interesa por la validez de los argumentos, no por si es razonable plantearlos. A la inversa, algunas personas son acusadas de «ilógicas» cuando en realidad los argumentos que proponen son válidos, aunque resulten absurdos por otros motivos.

A la lógica sólo le preocupa una cosa: que no haya forma de que las premisas sean ciertas sin que lo sea también la conclusión.

Fue Aristóteles (384-322 a.C.) quien intentó elaborar por primera vez una taxonomía sistemática de los argumentos válidos e inválidos. Se dio cuenta de que cualquier teoría que pretendiera alcanzar ese objetivo tendría que clasificar los argumentos en función de las pautas de razonamiento que revelan, es decir, en función de lo que llamamos su *forma*. Por ejemplo, una de las formas de razonamiento más famosas, que lleva el vistoso nombre de *modus ponendo*

ponens, o más brevemente *modus ponens*, no es más que:

p;

Si *p* entonces *q*;

Luego, *q*.

En este caso, *p* y *q* se refieren a cualquier fragmento de información, o proposición, que nosotros queramos. La forma del argumento será la misma tanto si hablamos de vacas como si hablamos de filósofos. Así pues, la lógica estudia las diferentes formas de presentar la información, no la información que se presenta en cada caso. Los argumentos concretos son instancias particulares de las formas, pero el lógico se interesa por su forma o estructura, del mismo modo que el matemático se interesa por la forma o estructura de los números, y no por si estamos contando plátanos o ganancias.

Queremos que nuestros razonamientos sean válidos. Ya hemos dicho lo que esto significa: queremos que no haya forma de que nuestra conclusión pueda ser falsa si las premisas son verdaderas, de modo que tenemos que estudiar si existe «alguna forma» de que cierto conjunto de cosas, las premisas, puedan ser verdaderas sin que otra cosa, la conclusión, también lo sea. Para investigar esta cuestión debemos elaborar una ciencia acerca de las distintas formas en que las cosas pueden ser verdaderas. Podremos hacerlo si nos fijamos en formas muy sencillas de formular la información.

Tablas de verdad

Los principios clásicos son, en primer lugar, que toda proposición (*p, q...*) posee tan sólo uno de los dos posibles valores de verdad: tiene que ser verdadera o falsa, y no puede ser ambas cosas. («Pero

supongamos que yo no concedo eso...» Paciencia.) El segundo principio es que los términos con los que trata la lógica — básicamente «y», «no», «o» y «si... entonces. ..»— pueden ser caracterizados a partir de su relación con los valores de verdad. («Pero supongamos que yo no concedo eso...» Paciencia.)

Así pues, consideremos «no- p ». No- p , que a menudo se escribe $\neg p$, constituye la negación de p : es lo que decimos cuando no estamos de acuerdo con p . Sea lo que sea lo que signifique p , de acuerdo con el primer principio, debe ser verdadera (V) o falsa (F). No puede ser ambas. ¿Cuál es la función del «no»? Simplemente, invierte el valor de verdad. Si p es verdadero, $\neg p$ es falso. Si p es falso, entonces $\neg p$ es verdadero. Eso es lo que hace el «no». Podemos resumir el resultado con una *tabla de verdad*:

p	$\neg p$
V	F
F	V

La tabla ofrece el resultado, en términos de verdad o falsedad, de cada asignación de un valor de verdad a los componentes (tal asignación recibe el nombre de *interpretación*). Se puede elaborar una tabla parecida para «y», con la única diferencia de que en este caso tendremos que considerar más combinaciones. Suponemos que «y» asocia dos proposiciones, cada una de las cuales puede ser verdadera o falsa. De modo que en este caso debemos considerar cuatro situaciones o interpretaciones posibles:

p	q	$p \& q$
V	V	V
V	F	F
F	V	F
F	F	F

En este caso obtenemos el valor de verdad de la fórmula entera, la

conjunción, a partir de la combinación de los valores de verdad de sus componentes: las cuatro interpretaciones posibles de la fórmula.

El hecho de que podamos elaborar estas tablas se expresa diciendo que tanto la conjunción como la negación *son funciones de verdad*, o que son *operadores de verdad*. La lógica proposicional elemental estudia las funciones de verdad. Aparte de «no» e «y», también incluye «o» (p o q , que se considera siempre verdadera excepto en el caso de que tanto p como q sean falsos); y una versión de «Si p entonces q », que se considera siempre verdadera excepto en el caso de que p sea verdadero y q falso. Si expresamos esto último como « $p \rightarrow q$ », la tabla de verdad sería:

p	q	$p \rightarrow q$
V	V	V
V	F	F
F	V	V
F	F	V

Todos estos términos también reciben el nombre de «operadores booleanos». Quienes estén familiarizados con las bases de datos o con las hojas de cálculo sabrán lo que es una búsqueda booleana, que se basa exactamente en la misma idea. Si realizamos una búsqueda de accesorios de más de cinco años conservados en un almacén de York, obtendremos un resultado cuando se encuentre un accesorio que cumpla ambas condiciones. Si realizamos una búsqueda de los clientes que no hayan pagado antes del 1 de diciembre, obtendremos el resultado inverso al de una búsqueda de los clientes que hayan pagado antes del 1 de diciembre. Si realizamos una búsqueda de los clientes que o bien hayan comprado una lavadora o bien una cortadora de césped, obtendremos como resultado todos los que compraron una y todos los que compraron otra.

Ya estamos en condiciones de comprender el sentido de algunas reglas de inferencia. Consideremos la regla de que a partir de « $p \ \& \ q$ » podemos derivar p (o igualmente q). Esta inferencia no puede llevarnos a ninguna falsedad, ya que la única interpretación (la línea superior) que da como verdadera a « $p \ \& \ q$ » también da como verdaderos a cada uno de los componentes. De modo que la regla es válida. También podemos comprender por qué el antes mencionado *modus ponendo ponens* es una regla válida. Contiene dos premisas, « p » y «Si p entonces q ». ¿Acaso podemos encontrar alguna interpretación (alguna «forma») en la que ambas sean verdaderas y en la que q no lo sea? No. Porque si p es verdadera, la única interpretación que permite que $p \rightarrow q$ sea verdadera también da como verdadera a q .

En este zoo se pueden encontrar algunos animales interesantes. Uno de ellos es la contradicción. Consideremos la siguiente fórmula:

$$p \ \& \ \neg p$$

Esto expresa una contradicción (el último y definitivo no-no). Y ahora podemos decirlo en un sentido preciso, ya que resulta fácil demostrar a partir de las dos tablas que tenemos que cualquiera que sea el valor de verdad de p , el valor de verdad de esta fórmula será F. No hay forma de que sea verdadera, porque cuando uno de los elementos de la conjunción es verdadero, el otro es falso: siempre hay un elemento falso. Y la tabla de verdad de la conjunción muestra que en tal caso la fórmula en conjunto es falsa.

Supongamos ahora que, para complicar todavía más la situación, lo negamos:

$$\neg (p \ \& \ \neg p)$$

Los paréntesis significan en este caso que el « \neg » que hay fuera niega todo lo que hay dentro. Funcionan igual que los paréntesis en

$3 \times (4 + 2)$, que hacen que el resultado sea 18, a diferencia de lo que sucedería si se tratara de $(3 \times 4) + 2$, cuyo resultado es 14. En lógica, los paréntesis tienen una importancia extraordinaria, igual que ocurre en aritmética: se pueden evitar muchas falacias en el razonamiento formal y también en el informal con sólo saber dónde van los paréntesis.

Esto se conoce como el campo de aplicación de las negaciones, las conjunciones y los demás operadores. En este ejemplo, la negación exterior se aplica sobre el resto de la fórmula, $\neg p \ \& \ \neg p$ tiene un sentido completamente distinto, puesto que se trata de la simple conjunción de p consigo mismo, lo cual, por cierto, es falso en caso de *que* p sea verdadero (decir que algo es falso dos veces no lo mejora en absoluto). Una de las fantásticas virtudes de la lógica formal es que hace que la gente sea sensible a las ambigüedades en el campo de aplicación, que se producen cuando no se sabe dónde van los paréntesis, o en otras palabras, cuando no se sabe cuál es el elemento que manda. Si no sabemos esto, no podemos saber de qué formas pueden ser verdaderas nuestras premisas y nuestras conclusiones y, por lo tanto, saber si hay alguna forma de que nuestras premisas sean ciertas sin que lo sea también la conclusión.

La nueva fórmula, $\neg (p \ \& \ \neg p)$, invierte la tabla de verdad de la contradicción anterior. Así pues, es verdadera sean cuales sean los valores de verdad de sus componentes. Esto es lo que se llama una *tautología*. Se trata de una noción importante. En lógica proposicional, si tenemos las premisas bla-bla-bla y la conclusión patatín-patatán, queremos que sea verdad que «Si bla-bla-bla entonces patatín-patatán» es una tautología. No existe ninguna interpretación (ninguna forma de asignar los valores de verdad) que haga que las premisas sean verdaderas y que la conclusión sea falsa. Cuando esto es así, el argumento es válido exactamente en el mismo sentido que hemos visto hasta ahora.

Hay una forma de descubrir si un argumento es válido o no que es lo bastante común como para merecer un nombre propio. Podemos saber si el argumento «Si bla-bla-bla entonces patatín-patatán» es válido añadiendo «no patatín-patatán» a «bla-bla-bla» y viendo si lo que obtenemos es una contradicción. Si obtenemos una contradicción, el argumento es válido. Esto es algo que se desprende directamente del principio de que no puede ser que las premisas sean verdaderas y la conclusión falsa. No existe ninguna interpretación ni ningún modelo que corresponda a esa situación. Lo impide la contradicción. Esto es lo que se llama «establecer un supuesto para llegar a una contradicción» o «establecer un supuesto para llegar a una reducción», que viene del nombre latino de este procedimiento: *reductio ad absurdum*, o reducción al absurdo. El argumento de san Anselmo que vimos en el capítulo 5 tenía esta forma.

En matemáticas, no existe solamente $2+2$, sino también $3 \times (2+2)$ y $((2+3) \times (2+2)) - 5$, y podríamos seguir eternamente. Con la información sucede lo mismo. Podemos producir fragmentos de información tan complejos como queramos mediante la aplicación reiterada de combinaciones de funciones de verdad, y seguiremos teniendo el control sobre las interpretaciones que los convierten en verdaderos o falsos.

Más fácil de lo que parece

Así pues, la lógica estudia la estructura de la información. Su objetivo es revelar esta estructura y, por lo tanto, revelar también cómo se derivan unas cosas a partir de otras: lo que se requiere para demostrar p y también aquello que se deriva de p , tanto en el caso de p como en el de cualquier otro complejo de información. La relación que existe entre la estructura y la demostración es simplemente la siguiente: la estructura nos permite comprobar si no hay forma de que las premisas sean verdaderas sin que lo sea también la

conclusión. Comprender la estructura de la información es comprender en qué casos puede ser verdadera.

Hasta ahora nos hemos ocupado del tipo de complejos de información que obtenemos a partir de la negación, la conjunción o la introducción de condicionales en las proposiciones. De acuerdo con lo que hemos visto hasta ahora, no veremos mucha diferencia entre las expresiones «Algunas personas son filósofos» y «Todas las personas son filósofos». No son más que dos ejemplos de proposiciones (p). Pero no somos capaces de entrar en el contenido de cada proposición y comprender en qué consiste la diferencia de significado.

La lógica moderna apareció en el momento en que un nuevo desarrollo permitió superar este obstáculo. Fue el lógico y matemático alemán Gottlob Frege (1848-1925) quien lo consiguió, con su *Begriffsschrift* («escritura de conceptos») de 1879. Consideremos el siguiente argumento: «Cualquier investigación termina en algún punto, luego existe un punto en el que termina toda investigación» (a veces se pretende que los fundacionalistas que tratamos en el capítulo 1 planteaban algo muy parecido a esto). Pero algo tiene que fallar en este argumento, pues resulta muy parecido a este otro: «Todo el mundo tiene una madre, luego hay alguien que es la madre de todo el mundo». O bien: «Todo el mundo se ata los cordones de los zapatos, luego hay alguien que le ata los cordones de los zapatos a todo el mundo». Cualquier persona que se enfrentara al problema antes de Frege podía ver perfectamente que algo fallaba, pero al no disponer de una adecuada comprensión de la forma en que este tipo de información está construida, no podía decir exactamente qué era.

La clave para comprender el logro de Frege es distinguir mentalmente dos clases muy distintas de información. La primera nos resulta muy familiar. Consiste en asignar un término a cada nombre o expresión que se refiere a una persona o una cosa en particular: «Bill es rico», «Tony sonríe», «Esto es una naranja». En estas expresiones

encontramos un término referido al sujeto (los nombres propios «Bill» y «Tony» y el pronombre demostrativo «esto»), y luego se dice de ellos lo que se desea resaltar: «es rico», «sonríe» o «es una naranja». Todos estos términos se refieren a condiciones que pueden cumplir las cosas. Reciben el nombre de «predicados»: las personas ricas satisfacen el predicado «es rico», y no las demás. En esto consiste una de las estructuras básicas de la información: la forma sujeto-predicado.

Ahora supongamos que hacemos algo inesperado: eliminamos el término que se refiere al sujeto. Lo único que nos queda es una oración incompleta o predicado: «es rico» y otras por el estilo. Podemos señalar más claramente el elemento que falta mediante una expresión llamada «variable», para la que se emplea habitualmente x, y, z, \dots , igual que se hace en álgebra, de modo que ahora tenemos « x es rico». En este caso ya no se trata de una oración que transmita algún tipo de información, porque no estamos diciendo que nadie en concreto sea rico. Es una oración en la que falta algo: en el lenguaje técnico de la lógica se conoce como «predicado» u «oración abierta».

Y ahora viene el truco de magia. Supongamos que le pido a alguien que aplique la oración abierta a cierto dominio, como por ejemplo una clase, o la ciudad de Nueva York, y que después me dé algún tipo de información. Esta persona podría reconstruir sin ningún problema un fragmento de información parecido al que teníamos al principio, con sólo atribuir un nombre a algún individuo particular y luego decir que es rico. Pero no es necesario que haga eso. Puede hacer otra cosa completamente distinta. Lo que puede hacer es decirme la cantidad de veces que se satisface el predicado. Y eso me lo puede decir sin necesidad de informarme de quiénes son las personas que lo satisfacen. Es como si apuntara con la « x » de la oración abierta a cada uno de los miembros del dominio, y anotara la frecuencia con la que obtiene un acierto. Supongamos que simboliza el predicado con la letra ϕ (la letra griega «fi»), y se pregunta sucesivamente acerca

de cada uno de los miembros del dominio: ¿Este es ϕ ? ¿Este es ϕ ? Luego viene y me cuenta lo que ha sucedido.

Tal vez la respuesta más sencilla que podría darme es que acertó al menos en un caso. Sería lo mismo que decir que «Alguno es ϕ ». O bien podría decirme que en algún caso no acertó: «Alguno no es ϕ ». Comparemos este caso con el de no conseguir ningún acierto: «Ninguno es ϕ ». O bien podríamos haber acertado en todos los casos: «Todos son ϕ ».

Para expresar que «Alguno es ϕ » emplearemos otro símbolo: el *cuantificador existencial*. Se escribe $(\exists x) \phi x$ (es irrelevante el hecho de que la variable se sitúe después del predicado, a diferencia de lo que sucede en inglés, en el que los predicados acostumbran a cerrar las oraciones y las abren los nombres y cosas por el estilo). Si no hemos conseguido ningún acierto, escribiremos $\neg(\exists x) \phi x$: ninguno es ϕ . Si en algún caso hemos encontrado un resultado distinto de un acierto, obtendremos una expresión muy distinta: $(\exists x)\neg(\phi x)$. Si no hemos encontrado ningún resultado distinto de un acierto, escribiremos $\neg(\exists x)\neg \phi x$. Esta fórmula significa que no hay ningún miembro que no sea ϕ , en otras palabras, que dentro del dominio todos son ϕ . Este último tipo de información tiene la suficiente importancia como para merecer un símbolo propio, el *cuantificador universal*, que se escribe $(\forall x) \phi x$: «Todos son ϕ ».

Leibniz pensaba que si pudiéramos encontrar un sistema de notación lo bastante lógico, terminarían todas las confusiones y las disputas, y las personas se sentarían a resolver sus diferencias mediante el cálculo. Aunque la invención de los cuantificadores no hizo realidad esta utopía, sí que dio algunos pasos espectaculares en esta dirección. Su verdadero poder sólo se descubre cuando llegamos a las cuantificaciones múltiples. Se trata de una forma de construir la información que emplea diversos cuantificadores. Cuando tenemos más de un cuantificador, utilizamos variables distintas (x, y, z, \dots) para

indicar a qué se refiere cada uno de ellos. Para ilustrar esta idea, podemos realizar un sencillo análisis del argumento inválido que veíamos antes: todo el mundo tiene una madre, luego alguien es la madre de todo el mundo. Si expresamos el hecho de que «x» es la madre de «y» diciendo que «xMy», podemos formular la primera cláusula como $(\forall y) (\exists x) xMy$. La segunda sería $(\exists x) (\forall y) xMy$. ¿Cuál es la diferencia entre ambas?

Comencemos con una oración que afirma que una persona es la madre de otra: Beth es la madre de Albert. Si eliminamos la referencia a Beth obtenemos la oración abierta xMa (donde «a» ocupa el lugar de Albert). Sabemos que este predicado se satisface en una ocasión (Beth lo satisface), de modo que sabemos que $(\exists x) xMa$. Alguien es la madre de Albert. Si ahora eliminamos la referencia a Albert, obtenemos $(\exists x) xMy$. De nuevo nos encontramos con una oración en la que faltan cosas, o abierta, en la que y ocupa el lugar de aquello que falta. Corresponde al predicado «tener a alguien por madre». Podemos aplicar este predicado al dominio y apuntar con la variable y a cada uno de los miembros sucesivamente: ¿tiene madre esta persona, tiene...? Si en cada caso obtuviéramos la respuesta «sí» (y eso es lo que sucedería), podríamos emplear el cuantificador universal: $(\forall y) (\exists x) xMy$: «Todos tienen una madre».

Examinemos ahora la segunda fórmula. Para llegar hasta ella, comenzaremos igual que antes con la afirmación de que Beth (b) es la madre de Albert. Pero ahora eliminaremos primero la referencia a Albert: bMy, y aplicaremos este predicado al dominio. Si fuera cierto que Beth es la madre de todo el mundo, el resultado debería ser $(\forall y) bMy$ (cada vez que apuntáramos la variable y hacia alguien, resultaría que Beth es su madre, lo cual evidentemente no sucede en el mundo real). Pero aunque este supuesto no sea cierto en el caso de Beth, puede que lo sea en el de alguna otra persona: en tal caso habría que eliminar la referencia a Beth, aplicar el predicado «ser la madre de

todos», o en otras palabras $(\forall y) xMy$, al dominio, y ver si existe realmente algún caso en que la respuesta sea sí. En tal caso podríamos escribir $(\exists x)(\forall y) xMy$. Pero lo que es realmente importante es que se trata de un procedimiento enteramente distinto. Nos da otro tipo de información (que en este caso es falsa en relación con el dominio de los seres humanos). La estructura de los cuantificadores muestra claramente la diferencia, ya que el modo en que éstos se relacionan entre sí revela la forma en que está construida la información.

En el mundo real, no hay nadie que sea la madre de todo el mundo. Antes de que comenzáramos a emplear los cuantificadores, esta afirmación podría haber resultado extraña, como si estuviéramos diciendo que la raza humana surgió de la nada. Habría sonado como una inquietante tesis metafísica. Pero ahora ha perdido punta. Sólo significa que $\neg(\exists x)(\forall y) xMy$, lo cual es una verdad muy sencilla. A menos que pretendamos estar usando la palabra «madre» en un sentido que incluya parentescos más remotos, en cuyo caso se podría decir que sí existe alguien, la Eva biológica, la primera hembra de *homo sapiens*, que sería efectivamente la madre de todo el mundo. Pero yo me inclinaría a pensar que se trata de un uso ilegítimo o metafórico de la palabra. Mi bisabuela no es mi madre en sentido literal.

Podemos precisar todavía más la información acerca de la cantidad de veces que se cumple una determinada condición en un dominio. Podríamos decir que hay exactamente una cosa que satisface la condición. Esto significa que cada vez que obtengamos un acierto, si seguimos apuntando con la variable hacia los demás objetos del dominio, cuando volvamos a obtener un acierto resultará ser el mismo. No es posible obtener dos aciertos distintos. Esta es la esencia de la famosa teoría de Russell de las descripciones definidas. Para que sea verdad que el único rey de Francia tiene barba, es necesario que exista alguien que sea el rey de Francia y que nadie

más sea el rey de Francia, y tiene que ser verdad que la persona que sea el rey de Francia tenga barba. En cualquier otro caso, es falso.

La estructura de los cuantificadores es tan sólo una de las cosas que debemos tener en cuenta, aunque tiene una importancia extraordinaria. En el lenguaje ordinario se producen frecuentemente ambigüedades que los cuantificadores resuelven con facilidad. Había una canción que decía: «Todas las chicas bonitas aman a un marinero». ¿Acaso hay un único marinero afortunado al que todas aman? ¿Acaso todas aman a un marinero, pero cada una ama a uno distinto? Si escogemos un marinero al azar, ¿hemos de pensar que todas le aman? Se trata de cosas muy distintas, que serán ciertas en circunstancias también muy distintas. Una ambigüedad parecida es responsable de algo así como treinta mil muertes al año en Estados Unidos. «Un Estado libre necesita una milicia bien regulada para su seguridad: no recortemos el derecho de las personas a poseer y llevar armas.» ¿Todas las personas? ¿O más bien las personas tomadas como un colectivo, igual que decimos: «El equipo debe tener un autobús»? Si los padres fundadores hubieran podido pensar en términos de cuantificadores, tal vez se habría evitado derramar mucha sangre.

El lenguaje y la lógica

La lógica estudia las estructuras de la información que acabamos de describir y, por supuesto, cualesquiera otras que se puedan encontrar, por complejas que sean. Pero hay otro aspecto en el trabajo del filósofo que consiste en determinar si la información contenida en el discurso ordinario se ajusta realmente a una u otra de estas estructuras. Y sorprende lo complicada que puede llegar a ser esta tarea.

Consideremos, por ejemplo, la diferencia entre las expresiones «Era pobre y honesta» y «Era pobre pero honesta». La primera

responde claramente a la forma « $p \ \& \ q$ ». Pero ¿qué se puede decir de la segunda? Ciertamente sugiere algo más, algo así como que resulta sorprendente o destacable que una persona pobre pueda ser honesta. Pero ¿acaso dice realmente eso? Una interpretación más sencilla sería que en un sentido estricto dice lo mismo que la primera, pero lo dice de una forma que insinúa o sugiere que la combinación resulta sorprendente o destacable. Tal vez en un sentido estricto no esté dando más que la información básica, pero lo hace de una forma que hace pensar en otras cosas (las cuales pueden ser abiertamente desagradables, como en este caso). De este modo, los filósofos del lenguaje se encuentran en la necesidad de distinguir lo que se dice o asevera en sentido estricto —la información que contiene la expresión lingüística, llamada su *condición de verdad*— de aquello que se sugiere o se implica, no como una consecuencia lógica estricta, sino a través de la forma en que se dicen las cosas, llamada la *implicatura*.

El lenguaje es un instrumento tan flexible y sutil que ofrece un margen prácticamente ilimitado para manipular la información y, por lo tanto, las implicaturas. El mero hecho de no decir algo puede tener sonoras consecuencias:

«¿Cuál es tu opinión acerca del nuevo profesor?»

«Me han dicho que es famoso por sus tomates.»

En este caso, lo que se ha dicho guarda poca o nula relación con la competencia profesional del nuevo profesor. Pero el hecho de que no se dé más respuesta que ésta revela de forma incuestionable que en su opinión el profesor no es muy bueno. La elección de los términos que se emplean también comporta sus propias implicaturas: consideremos la diferencia entre las expresiones:

John es el hermano de Fred.

Fred tiene un compañero masculino en la familia, John.

En este caso, la segunda forma de presentar una información que

viene a ser la misma parece apuntar en alguna otra dirección —tal vez algún siniestro caso de desviación psicoanalítica—. El orden en el que se dicen las cosas también conlleva implicaturas acerca del orden en que se produjeron los acontecimientos. Si explicásemos la vida de una niña que aprendió a leer y luego escribió poesía diciendo que escribió poesía y aprendió a leer estaríamos induciendo a confusiones, por muy cierto que fuera lo que habíamos dicho.¹

La rama del estudio del lenguaje que se ocupa de la forma en que se generan las implicaturas se llama *pragmática*, mientras que la *semántica* se encarga de la estructura de la información.

Pensemos en la temible pregunta que emplean los abogados para desmontar el testimonio de los hombres casados: «¿Ha dejado usted de pegar a su esposa? Responda sí o no». El testigo no puede responder «sí» sin admitir implícitamente que en alguna ocasión lo hizo; tampoco puede responder «no» sin dar toda la impresión de que todavía lo hace. De modo que el truco deja al testigo desconcertado, que es precisamente lo que buscaba. ¿Se nos ocurre algo mejor que hacer? Pues bien, supongamos que interpretamos la expresión: «X dejó de hacer Y» como una conjunción: «X alguna vez hizo Y & X ya no hace Y». Esto explica por qué es malo responder «sí» al abogado: se desprende de ello que en alguna ocasión sí pegó a su esposa. Por otro lado, decir «no» resulta interesante. Si nos fijamos en la tabla de verdad de la conjunción nos damos cuenta de que puede ser falsa en tres casos distintos: que p sea verdadera y q falsa; que p sea falsa y q verdadera, y que ambas sean falsas. Y en los tres casos la negación de la conjunción es verdadera (la negación invierte el valor de verdad). En nuestro caso, es vital para el marido inocente que deje claro que se encuentra en el segundo caso: es falso que alguna vez pegara a su esposa, y verdadero que no lo hace actualmente. El problema consiste en que la palabra «no» es insuficiente para establecer a cuál de los tres casos se refiere y existe el riesgo de que el jurado lo interprete en el sentido de que no ha dejado de pegarla

porque todavía lo hace (es verdad que alguna vez lo hizo y también que lo hace ahora, luego es falso que dejara de hacerlo).

El testigo inocente necesita emplear más palabras para especificar cuál de las combinaciones es la que describe su situación, de modo que no puede quedarse con la escueta respuesta: «No» (por muy cierta que sea). Lo que debería hacer nuestro testigo es responder de una tirada: «No he dejado de hacerlo porque nunca lo hice», u otras palabras que vengan a decir lo mismo. De acuerdo con nuestro análisis de la pregunta del abogado, se puede decir que éste «presupone» que el testigo pegó en alguna ocasión a su esposa, pero sólo en un sentido pragmático, es decir, en el sentido de que cuando alguien hace esta pregunta lo normal es que lo esté dando por sentado. Pensar significa, en buena medida, revelar las presuposiciones que se ocultan detrás de las preguntas y las opiniones.

Algunas presuposiciones llegan incluso a plantear dudas acerca de la validez de nuestra interpretación de los términos «y», «no», «o», y sobre todo «si... entonces...», como funciones de verdad que se pueden describir adecuadamente en las tablas de verdad. A veces parece que su uso sea más complicado. Por ejemplo, imaginemos que Fred está invitado a una fiesta, pero que finalmente no asiste a ella. Supongamos que hay dos asesinos que pretenden seguir los movimientos de Fred. Uno dice: «Si Fred va a la fiesta, irá en taxi». El otro dice: «Si Fred va a la fiesta, irá en elefante». Intuitivamente, podemos decir que tan sólo una de las dos suposiciones puede aspirar a ser verdadera: la primera, por lo menos en Occidente. Pero si lo analizamos («Fred va a la fiesta→...») veremos que ambas son verdaderas. La razón es que «Fred va a la fiesta» es falso, y en tal caso la tabla de verdad de → indica que será verdadera sea cual sea el valor de verdad de la otra proposición. La posibilidad de que este hecho revele que el uso del condicional «Si... entonces» en el inglés es distinto de la función de verdad → ha suscitado un encendido

debate entre los filósofos. En los últimos años, dicho debate se ha relajado un poco, en la medida en que se acepta que → resume en cualquier caso la idea fundamental, y que todo lo demás se puede explicar desde el punto de vista semántico o pragmático.

Antes de terminar este breve resumen de la lógica formal, deberíamos detenernos un momento a considerar cierta clase de reacciones que a veces provoca. La gente piensa en ocasiones que la lógica es coercitiva («masculina») o que supone privilegiar cierta clase de «pensamiento lineal» que se opondría al «pensamiento lateral». Ambas acusaciones son completamente injustas. La lógica formal es demasiado modesta como para merecerlas.

En primer lugar, ¿qué significa esta acusación de que la lógica es coercitiva? La lógica formal nos permite determinar si un determinado conjunto de proposiciones implica una contradicción. Por otro lado, considera que las contradicciones son falsas. La mayoría de nosotros preferimos no creer en un conjunto de proposiciones que nos lleve a conclusiones falsas, ya que es importante para nosotros que nuestras creencias sean verdaderas. Si alguien piensa de otro modo, es muy posible que intentemos convencerle de nuestras ideas. Pero cuando hacemos eso, no estamos actuando como lógicos. La tarea del lógico formal termina en cuanto obtiene el resultado.

Tal vez alguien vea la coerción en el principio que hemos mencionado al comienzo: que toda proposición debe ser verdadera o falsa, y que no puede ser ambas. Tal vez deberíamos ensayar principios más elaborados: por ejemplo, podríamos admitir proposiciones vagas, que fueran verdaderas sólo hasta cierto punto, o proposiciones que no fueran ni verdaderas ni falsas, sino que pertenecieran a un tercer grupo. Todo esto está muy bien: son ideas respetables y existen lógicas alternativas que las desarrollan. Pero vale la pena indicar que, por varias razones, acaban convirtiéndose en modelos extravagantes e incómodos. En la mayoría de los casos se agradece el sencillo y razonable principio de los «dos valores».

El tercer motivo de este sentimiento de coerción nos lleva a cuestiones más amplias. Cuando alguien expone una serie de opiniones, o plantea algún razonamiento, puede resultar torpe y coercitivo insistir en que tienen tal o cual forma y que son, por lo tanto, contradictorios, o inválidos. Es muy posible que esta actitud sea insensible a los otros factores que hemos mencionado: presuposiciones, premisas ocultas y otros por el estilo. Pero esto no es culpa de la lógica, sino de la respuesta tan poco caritativa que se ha dado. Por su propia naturaleza, la lógica es indiferente a todo esto, incluso a las afirmaciones que parecen expresar contradicciones explícitas. En el cuento de Chejov titulado *La señora del perrito*, Anna Sergeyevna le dice a su marido cada tanto tiempo que se va a Moscú a visitar a un médico, «y el marido la creía y no la creía». La lógica formal no nos pide que arrojemos al fuego el libro de Chejov a causa de esta sangrante contradicción. Sabemos perfectamente que Chejov está sugiriendo otra cosa, a saber, que el marido cree en ella a medias, o que alternativamente confía y desconfía. Y es precisamente la contradicción explícita lo que nos hace pensar en otras posibles interpretaciones. ¿Cómo podemos responder a la acusación de que la lógica privilegia el «pensamiento lineal»? Pues bien, esta crítica tampoco tiene sentido. La lógica formal no dirige los pensamientos de nadie, del mismo modo que las matemáticas no nos dicen lo que debemos contar o medir. Le resulta magníficamente indiferente si las proposiciones proceden de la especulación, la imaginación, el mero capricho, la escrupulosa ciencia o lo que sea. Todo lo que nos dice es si existe alguna forma de que las proposiciones que forman parte de un conjunto, con independencia de cómo se haya llegado hasta él, pueden ser verdaderas dentro de este conjunto. Pero eso puede llegar a tener un valor incalculable.

Razonamientos plausibles

La lógica formal es de gran ayuda a la hora de evitar las contradicciones. De modo parecido, nos ayuda a saber qué es lo que podemos deducir a partir de un determinado conjunto de premisas. Pero antes tenemos que llegar a estas premisas. Nuestros razonamientos no sirven tan sólo para deducir cosas a partir de una información dada, sino también para ampliar nuestras creencias, o aquello que consideramos información. De modo que algunos de los razonamientos más interesantes que realizamos en la vida cotidiana no pretenden ser válidos según los criterios que hemos descrito hasta ahora. Pretenden ser plausibles o razonables, más que incontestables. Es posible que un argumento de este estilo tenga unas premisas verdaderas y una conclusión falsa, aunque no es probable que ocurra.

Sin embargo, podemos llevar un poco más lejos la aplicación de algunos de los principios que hemos visto, incluso en el caso de los razonamientos plausibles. Por ejemplo, ¿por qué habría de ser ingenuo confiar en que mi apuesta en la ruleta va a ser la ganadora? Pues porque la única información de la que dispongo es que he apostado por x , y que la mayoría de posiciones en las que puede terminar la bola al detenerse la ruleta no dan a x como ganador. En este tipo de casos lo que importa es el margen de probabilidades, y el hecho de que nuestro razonamiento se pueda considerar plausible o no depende de si podemos demostrar que la mayoría de posibilidades que permite la evidencia de la que disponemos hacen que nuestra conclusión sea verdadera. En el ejemplo de la ruleta, la evidencia señala que la mayoría de resultados posibles falsean la conclusión de que x es la ganadora.

La ruleta y otros juegos de azar son pequeños modelos diseñados especialmente para que podamos saber cuáles son las posibilidades y podamos calcular las probabilidades de que ocurran. Cuando sacamos una carta tenemos cincuenta y dos resultados posibles, y si la sacamos de una baraja que ha sido mezclada recientemente y de forma honesta, cada una de estas posibilidades tiene las mismas

probabilidades de ocurrir. El razonamiento probabilístico puede trabajar entonces sin problemas: podemos calcular, por ejemplo, la probabilidad de que la mayoría de las veces que saquemos siete cartas al menos dos de ellas sean figuras, o cualquier cosa por el estilo. El razonamiento probabilístico consiste precisamente en calcular el margen de posibilidades que abre la especificación y comprobar en qué proporción se produce cierto resultado.

¿En qué nos basamos para asignar probabilidades en el mundo real? Supongamos que planteamos nuestra posición del siguiente modo: a lo largo de nuestra vida, obtenemos una cierta experiencia acerca de cómo son las cosas. Dentro del marco de nuestra experiencia parece que es posible hacer ciertas generalizaciones: la hierba es verde, el cielo azul, el agua refresca, el chocolate alimenta; de modo que tomamos esta experiencia como modelo acerca de cómo deben de ser las cosas en un marco más amplio de espacio y tiempo. No tengo ninguna experiencia directa de si el chocolate alimentaba en el siglo XVIII, pero supongo que era así; no tengo ninguna experiencia directa de si lo hará mañana, pero supongo que seguirá siendo así. Nuestras creencias y nuestra confianza se extienden más allá del limitado círculo de los hechos que se producen en el campo de nuestra experiencia inmediata.

Hume planteó el problema del siguiente modo:

Con respecto a la *experiencia* pasada, sólo puede aceptarse que da información directa y cierta de los objetos de conocimiento y exactamente de aquel período de tiempo abarcado por su acto de conocimiento. Pero por qué esta experiencia debe extenderse a momentos futuros y a otros objetos, que, por lo que sabemos, puede ser que sólo en apariencia sean semejantes, ésta es la cuestión en la que deseo insistir. El pan que en otra ocasión comí, que me nutrió, es decir, un cuerpo con determinadas cualidades, estaba en aquel momento dotado con determinados poderes secretos. Pero ¿se sigue de esto que otro trozo distinto de pan también ha de nutrirme en otro momento y

que las mismas cualidades sensibles siempre han de estar acompañadas por los mismos poderes secretos? De ningún modo parece la conclusión necesaria. Por lo menos ha de reconocerse que aquí hay una conclusión alcanzada por la mente, que se ha dado un paso, un proceso de pensamiento y una inferencia que requiere explicación. Las proposiciones siguientes distan mucho de ser las mismas: *He encontrado que a tal objeto ha correspondido siempre tal efecto y preveo que otros objetos, que en apariencia son similares, serán acompañados por efectos similares.* Aceptaré, si se desea, que una proposición puede correctamente inferirse de la otra. Sé que, de hecho, siempre se infiere. Pero si se insiste en que la inferencia es realizada por medio de una cadena de razonamientos, deseo que se presente aquel razonamiento.²

La experiencia no va más allá de ciertos límites en el espacio y en el tiempo. En concreto, el conjunto de nuestra experiencia se limita al pasado y al presente. Nuestras inferencias acerca del futuro no son más que eso, inferencias, y a Hume le interesa saber cuál es la «cadena de razonamiento» que empleamos para realizarlas.

La inferencia que parte de lo que se puede afirmar acerca de una región limitada del espacio y del tiempo para deducir lo que se puede afirmar acerca de regiones distintas del espacio y del tiempo se llama inferencia inductiva. La cuestión que preocupa a Hume en este punto se conoce habitualmente como el problema de la inducción.

La lotería del arpa dorada

Hagamos un poco de ciencia ficción. Vosotros sois espíritus sin cuerpo que viven en una especie de cielo. Yo soy Dios, y os informo de que os voy a dar un cuerpo para que podáis vivir una vida en el universo físico que he creado especialmente para vosotros: la Tierra. Al final de vuestra vida en este universo, volveréis al cielo. A diferencia de lo que ocurre en el caso de la vida humana, todos viviréis el mismo período de tiempo: nueve actos, digamos.

Este boleto corresponde a la posibilidad de que el cielo comience por

naranja
amarillo
verde
azul
violeta

Color	Frecuencia
naranja	6
amarillo	4
verde	3
azul	2
violeta	1

ser azul, se vuelva verde, y pase luego al amarillo y al naranja antes de oscurecerse de nuevo hasta llegar al azul e incluso al violeta.

Digamos que la persona que posee este boleto se llama Ondulado.

Algunos de vosotros (seis en total) obtenéis boletos de este tipo:

rojo									
naranja									
amarillo									
verde									
azul	X	X	X	X	X	X	X	X	X
violeta									
Tiempo	1	2	3	4	5	6	7	8	9

Este boleto se podría llamar Recto.

Si cada uno de vosotros va a tener un solo boleto, es necesario que seáis 6^º espíritus (un número muy elevado) para que entre todos reunáis todas las posibles combinaciones de colores. Y de acuerdo con eso la probabilidad de ganar es de 1/6^º, es decir, muy pequeña.

Hume insiste en que cuando nos encontramos en una situación como ésta no podemos saber nada al principio. No tenemos ningún conocimiento *a priori* de cuál va a ser el boleto ganador. Antes de

llegar a eso, y mientras todavía os dedicáis a discutir animadamente acerca de los boletos, no hay motivo alguno para preferir uno u otro. Por lo que sabéis, Dios podría dar preferencia a las ondas o a las rectas. O bien podría dar preferencia al Torcido:

~v~									
naranja									
amarillo						x	x	x	x
verde									
azul	x	x	x	x	x				
violeta									
Tiempo	1	2	3	4	5	6	7	8	9

El cielo de mediodía de los primeros cinco actos sería azul, luego se volvería amarillo y se mantendría así durante el resto de los actos. Así que mientras todavía estáis en el cielo, antes de tener la menor experiencia acerca del mundo al que os va a enviar Dios, ningún boleto tiene mejores probabilidades que otro.

Pues bien, ahora bajemos a la Tierra.

Inmediatamente, $5/6$ de vosotros podéis tirar vuestro boleto. Cualquier boleto que no indique azul en la primera casilla ha perdido. Y del mismo modo, el primer día de cada nuevo acto, $5/6$ de los supervivientes pueden tirar su boleto, hasta que al comienzo del noveno acto sólo quedan seis. Y un día más tarde tan sólo queda el ganador.

Ahora vayamos al final del quinto acto. Tanto Recto como Torcido han tenido buena fortuna. Han visto cómo sus competidores han ido cayendo en las cinco ocasiones anteriores. En realidad, el número de participantes en la lotería ha bajado de 6^9 a 6^4 , con lo que sus probabilidades de ganar han aumentado en la misma medida.

Pero supongamos que se ponen a discutir entre ellos. Supongamos que Recto trata de convencer a Torcido de que su boleto tiene muchas más probabilidades de ganar, por lo que sólo estará dispuesto a hacer un trueque a cambio de un altísimo precio. Es muy probable que todos estuviéramos de acuerdo con Recto. Pero supongamos que Torcido se resiste a aceptarlo, diciendo que lo que haya podido suceder hasta el momento no es motivo para apostar a favor de Recto antes que a favor suyo. ¿De qué argumentos dispone

cada uno?

Cada uno de ellos puede destacar su propio historial de éxitos, aunque será el mismo en ambos casos. Ambos han obtenido sus cinco aciertos. Y no hay nada más que decir. Después de todo, ninguno de ellos puede ver el futuro. Igual que nosotros, están atrapados en el tiempo y no pueden salir de él.

Lo que Recto desearía encontrar es un argumento en favor de la uniformidad de la naturaleza. En otras palabras, un argumento que dijera que puesto que Dios ha comenzado con un cielo azul, y lo ha mantenido hasta ahora, es probable que lo siga manteniendo. A lo que Torcido podría responder que Dios ha comenzado con un cielo de tipo Torcido, y que por el mismo motivo va a mantenerlo.

Recto persigue el mismo argumento que Hume afirma que no se puede encontrar. Sin embargo, tal como he señalado antes, en el fondo todos estamos de acuerdo con Recto. ¿Qué hay de malo en el argumento de que si la naturaleza ha sido uniforme hasta el momento, es probable que lo siga siendo?

Es imposible, por tanto, que cualquier argumento de la experiencia pueda demostrar esta semejanza del pasado con el futuro, puesto que todos los argumentos están fundados sobre la suposición de aquella semejanza.³

Por supuesto, Hume es consciente de que aprendemos a partir de la experiencia, y de que confiamos en la uniformidad de la naturaleza. Opina que es una propensión natural que compartimos con los animales. Lo único que dice es que no es más que eso: una tendencia natural. Se trata de un hábito o una costumbre que no puede reclamar ningún derecho especial ante la razón. Cuando razonamos de forma inductiva, existe una posibilidad de que nuestras premisas sean verdaderas y la conclusión falsa. La naturaleza puede cambiar. En realidad, tal posibilidad puede materializarse de muchas formas, puesto que la naturaleza puede cambiar de muchas formas distintas.

Pensar de este modo no nos lleva a ninguna contradicción. Y según parece no podemos ni siquiera argumentar que tales cambios son improbables. El único motivo de que nos lo parezcan es que no se han producido nunca en el marco de nuestra experiencia: Sin embargo, considerar que nuestra experiencia es representativa presupone, tanto en este sentido como en muchos otros, que la naturaleza es uniforme. Parece que somos capaces de levantar un puente entre el pasado y el futuro, pero que no podemos asegurar que sea fiable.

Cuestión de probabilidades

A continuación vamos a examinar un problema relativamente alejado del problema de la inducción, pero que introduce una herramienta increíblemente útil para tratar un buen número de cuestiones. Se trata de un problema que la mayoría de personas plantean mal.

Supongamos que decides hacerte una prueba para saber si padeces cierta enfermedad. Supongamos que se trata de una enfermedad bastante inusual entre la población: sólo la tiene una de cada mil personas. Sin embargo, encuentras a un doctor que afirma disponer de un buen test para detectarla. En realidad, su fiabilidad es superior al 99%. Al ser informado de esto, decides someterte al test. Y entonces —¡horror!— resulta que das positivo. Téngase en cuenta que el test posee una fiabilidad superior al 99%. ¿Hasta qué punto es mala tu situación, o en otras palabras, qué probabilidades tienes de padecer la enfermedad?

La mayor parte de la gente diría: es terrible, tienes prácticamente la certeza de padecer la enfermedad.

Pero supongamos que eres experto en probabilidades y que le pides algunas precisiones al doctor acerca de este 99 % de fiabilidad. Supongamos que obtienes la siguiente información:

- 1) Si tienes la enfermedad, el test dirá que la tienes.

2) Algunas veces, aunque muy raramente, el test da «falsos positivos». El test sólo se equivoca al indicar que alguien tiene la enfermedad en muy pocos casos (alrededor del 1%).

El significado de aquel 99 % de fiabilidad son estos dos principios. Podrías pensar que sigues teniendo prácticamente la certeza de padecer la enfermedad. Y sin embargo esta impresión es completamente errónea. De acuerdo con los datos de que dispones, la probabilidad de que tengas la enfermedad es algo menor al 10%.

¿Cómo es posible? Pues bien, supongamos que 1.000 personas se someten al test. De acuerdo con la incidencia general de la enfermedad (la «probabilidad de base»), se puede esperar que una de estas personas tenga la enfermedad. El test así lo indicará. También indicará que el 1 % de las demás personas que han pasado el test, es decir, aproximadamente diez personas, tienen la enfermedad. De modo que se puede esperar que 11 personas den positivo, de las cuales tan sólo una tendrá la enfermedad. Ciertamente se trataba de una mala noticia —has pasado de una probabilidad de 1 entre 1.000 de tener la enfermedad a otra de 1 entre 11—, pero sigue siendo mucho más probable que te encuentres en perfecto estado de salud. El clásico error que cometemos en estos casos se conoce como la falacia de ignorar la probabilidad de base.

¿Cuál es la mejor forma de plantear la cuestión de las probabilidades en un caso como éste?

Lo primero que deberíamos encontrar es una fórmula que exprese la probabilidad de una cosa dada otra. Supongamos que nos preguntamos cuál es la probabilidad de que a) una persona tomada al azar dentro de una clase lleve unos vaqueros Levi. Tal vez del 20%. ¿Y cuál es la probabilidad de que b) una persona tomada al azar lleve una chaqueta Levi? Quizás del 20 % también. Así pues, ¿cuál es la probabilidad de que una persona tomada al azar lleve al mismo tiempo los vaqueros y la chaqueta? Tal vez lo primero que

pensaríamos es que $20/100 \times 20/100 = 4\%$. Pero estaríamos cometiendo un error, ya que los dos sucesos no tienen por qué ser necesariamente independientes. Es decir, la probabilidad de que alguien lleve la chaqueta seguramente varía en función de si lleva o no los pantalones. Tal vez todas las personas que llevan los pantalones llevan también la chaqueta, y al revés. En tal caso, la probabilidad de que $a)$ y $b)$ sean verdaderos al mismo tiempo en relación con una persona tomada al azar es también del 20%. O tal vez los gurús de la moda digan que no se deben llevar nunca los dos juntos, en cuyo caso la probabilidad de que $a)$ y $b)$ fueran verdaderos al mismo tiempo bien podría ser cero.

Para expresar adecuadamente este hecho necesitamos una fórmula que nos dé la probabilidad de que alguien lleve la chaqueta dado que lleva los vaqueros. La probabilidad de $a)$ dado $b)$ se expresa como $\text{Prob}(a/b)$. La probabilidad de $b)$ dado $a)$ sería $\text{Prob}(b/a)$. Así pues, la forma correcta de expresarlo es la siguiente:

$$\text{Prob}(a \& b) = \text{Prob}(a) \times \text{Prob}(b/a)$$

O bien:

$$\text{Prob}(a \& b) = \text{Prob}(b) \times \text{Prob}(a/b)$$

La primera ecuación afirma que la probabilidad de llevar los vaqueros y la chaqueta = la probabilidad de llevar los vaqueros x la probabilidad de llevar la chaqueta dado que llevamos los vaqueros. Esto último recibe el nombre de «probabilidad condicional». La segunda dice también que es igual a la probabilidad de llevar la chaqueta x la probabilidad de llevar los vaqueros dado que llevamos la chaqueta. Las dos ecuaciones tienen que ser equivalentes, por simetría (puesto que $a \& b$ es la misma proposición que $b \& a$).

Un clérigo inglés llamado Thomas Bayes (1702-1761) se planteó seriamente este resultado. Puesto que cada una de las ecuaciones es igual a $\text{Prob}(a \& b)$, tienen que ser equivalentes entre sí:

$$\text{Prob } (a) \times \text{Prob } (b/a) = \text{Prob } (b) \times \text{Prob } (a/b)$$

De modo que podemos dar una fórmula para la probabilidad de b dado a :

Esta ecuación de aspecto tan intimidatorio no es más que una versión de lo que se conoce como el teorema de Bayes.

El resultado que hemos obtenido se aplica del siguiente modo:

$$\text{Prob } (h/e) = \frac{\text{Prob } (h) \times \text{Prob } (e/h)}{\text{Prob } (e)}$$

supongamos ahora que tenemos una hipótesis y cierta evidencia en su favor. Nos interesa saber la probabilidad de la hipótesis h , dada la evidencia e . Podemos expresarlo como $\text{Prob } (h/e)$. Es lo que se conoce como la probabilidad *posterior* de una hipótesis, es decir, su probabilidad después de conocer la evidencia. En tal caso el teorema dice lo siguiente:

$$\text{Prob } (h/e) = \frac{\text{Prob } (h) \times \text{Prob } (e/h)}{\text{Prob } (e)}$$

Esto nos lleva a considerar que la probabilidad posterior depende de tres factores distintos:

$\text{Prob } (h)$. Se conoce como la probabilidad previa o antecedente de h .

$\text{Prob } (e/h)$. Se trata de la probabilidad de la evidencia e dado h . Sirve para medir el grado de *concordancia* que hay entre la hipótesis y la evidencia.

$\text{Prob } (e)$. Se trata de la probabilidad previa o antecedente de la propia evidencia.

Intuitivamente, podemos plantear la cuestión del siguiente modo: tenemos tres factores. Primero, ¿cuál es la probabilidad de partida de la hipótesis? Segundo, ¿hasta qué punto concuerda la evidencia con la hipótesis? Tercero, ¿cuál es la probabilidad de partida de la evidencia?

A menudo resulta útil plantearse esta última cuestión en términos de las diferentes explicaciones posibles acerca de cómo se ha llegado a producir la evidencia. Se trata de un factor que aumenta en función del número de explicaciones alternativas y razonablemente probables que se pueden plantear acerca de la evidencia. Y cuanto mayor es

este factor, menor es la probabilidad de la hipótesis planteada a partir de la evidencia. Tiene demasiados competidores. Así pues, en la práctica, lo que hace el factor que aparece debajo del signo de división es medir la cantidad de hipótesis alternativas capaces de explicar la evidencia, y la probabilidad que les corresponde. Podemos reconocer de forma intuitiva la importancia de este factor. Cuando la prostituta Mandy Rice-Davies supo que cierto miembro de la aristocracia negaba que hubiera tenido una aventura con ella, su respuesta fue: «Bueno, ¿es lo que tenía que decir, o no?». Lo que hizo fue recordarle a la gente que la probabilidad antecedente de este tipo de declaraciones era alta, con independencia de cuál fuera la hipótesis verdadera, lo que cuestionaba su valor como evidencia. Era perfectamente predecible que el aristócrata diría precisamente eso, fueran cuales fueran sus relaciones. De modo que el hecho de que hiciera esa declaración no tenía ningún valor como evidencia.

Lo ideal sería que la hipótesis tuviera una elevada probabilidad antecedente; que la evidencia fuera justo la que cabría esperar a partir de la hipótesis; y que no hubiera muchas otras explicaciones probables acerca de cómo podría haberse producido la evidencia.

En el caso de la enfermedad, el teorema de Bayes da la mayor importancia a la probabilidad de base: es la probabilidad antecedente de que tengamos la enfermedad, es decir, 1 entre 1.000. El siguiente factor, el grado de concordancia entre el resultado del test y la hipótesis de que tenemos la enfermedad, es excelente: en realidad es 1, ya que cuando tenemos la enfermedad el test siempre lo dice. Pero en la línea de abajo tenemos el número de explicaciones alternativas de la evidencia. Si lo calculamos de modo informal, tendríamos la probabilidad de 1 entre 1.000 de que el resultado sea correcto más la probabilidad de 10 entre 1.000 de que el positivo sea falso. El resultado es la probabilidad total, de acuerdo con la evidencia, que es (aproximadamente) de 1 entre 11.

También disponemos ahora de una buena forma de representar el

$$\text{Prob}(R/E) = \frac{\text{Prob}(R) \times \text{Prob}(E/R)}{\text{Prob}(E)}$$

impasse en el que se encontraban Recto y Torcido a propósito de la lotería del arpa dorada. Supongamos que la combinación completa de nueve posibilidades —todas azul— del boleto de Recto es R. Y supongamos también que E es la parte de ella que hemos podido comprobar mediante la experiencia: los cinco resultados que hemos visto hasta ahora. En tal caso

La probabilidad previa o antecedente de R es $1/6^9$. El segundo factor, en cambio, es bueno. Si R describiera realmente la forma en que se van desarrollar los hechos, la evidencia E, a saber, los cinco primeros resultados, son exactamente lo que cabría esperar. En realidad, su probabilidad es de 1, de acuerdo con R. ¿Y la probabilidad previa de E? Se trata de que haya cinco resultados de azul, lo cual, dado que azul es una posibilidad entre seis que hay en juego, da $1/6^5$. Si realizamos el cálculo, obtenemos que $\text{Prob}(R/E)$ es $1/6^4$, que es exactamente el resultado al que habíamos llegado antes de forma intuitiva.

El problema es que en el caso del boleto de Torcido (T) tenemos exactamente la misma fórmula y obtenemos exactamente el mismo resultado. Resulta fácil ver que la evidencia tiene una probabilidad de 1, dado K, y que la probabilidad previa de la evidencia es la misma en ambos casos.

Es importante señalar que el problema no consiste en «demostrar» que R va a ganar o que T va a perder, sino simplemente en encontrar alguna buena razón para esperar que ocurra R en lugar de T. Se trata de comparar probabilidades. La postura de Hume es que ni siquiera eso podemos hacer en beneficio de R. La razón mantiene un completo silencio ante las dos posibilidades. Y si tenemos en cuenta el teorema de Bayes, parece ser que tiene razón. El debate entre Recto y Torcido está tan empatado como el primer día. En realidad, si no había razón para preferir el boleto R al boleto T en el Cielo, *a priori*, no hay razón tampoco para preferirlo una vez que hemos tenido alguna experiencia del mundo. O eso parece.

Ahora podríamos replantearnos una serie de cuestiones desde la

perspectiva del teorema de Bayes: la opción zombi, el argumento del diseño, la posibilidad de que un buen Dios creara o permitiera el mal, y sobre todo la discusión acerca de los milagros. Se trata de una herramienta de enorme importancia. Nos protege de peligrosas falacias —ignorar la probabilidad de base, ignorar la posibilidad de que haya pruebas falsas—, que reaparecen una y otra vez cuando nos ponemos a reflexionar.

Por supuesto, a menudo resulta difícil o incluso imposible cuantificar con algún tipo de precisión las probabilidades «previas». Aunque es importante señalar que esta cuestión no tiene por qué ser tan importante como podría parecer. Hay dos factores que vienen a mitigar el problema. En primer lugar, incluso si lo que hacemos es asignar un cierto margen de valores a cada factor, puede ser que todos los resultados que obtengamos sean suficientemente parecidos. Y en segundo lugar, puede ser que si disponemos de la suficiente evidencia, las variaciones en la probabilidad previa se conviertan en irrelevantes. Por muy diferentes que sean las actitudes previas de una serie de investigadores hacia $\text{Prob}(h)$, es posible que terminen por asignar valores igual de altos a $\text{Prob}(h/e)$ en caso de que (e) llegue a tener el suficiente peso.

Por el interés que tienen, vale la pena señalar que existen métodos bastante ortodoxos de inferencia estadística que intentan seguir un camino distinto al de las ideas de Bayes. Muchas investigaciones científicas se dan por satisfechas con sólo demostrar que cierto resultado se produce de forma azarosa tan sólo en un porcentaje muy limitado de ocasiones (menos del 5%, o por ejemplo menos del 1%). Pero luego infieren que es probable que el resultado no se deba al azar, es decir, que habría que considerar la presencia de algún elemento causal significativo o de algún tipo de correlación. Esta forma de razonar, que cuenta con una aceptación mayoritaria, resulta, sin embargo, altamente cuestionable. Bayes señala el motivo: si la probabilidad antecedente de que cierto resultado se deba a algo

distinto que el azar es muy, muy baja, ni siquiera los resultados más improbables van a poder girar la balanza. Si pongo la mano en una bolsa que ha sido previamente agitada, saco siete letras de Scrabble, las pongo boca abajo sobre la mesa, las ordeno para que formen una línea y luego les doy la vuelta, el resultado que obtendré (PQAERTU, por ejemplo) será ciertamente de lo más improbable. Podría ser que hiciera lo mismo durante cien años y que no se volviera a repetir. Pero a pesar de todo fue una cuestión de azar. Si seguimos este procedimiento, cualquier resultado que obtengamos será de lo más improbable, y no deberíamos sentirnos con derecho a inferir que alguna otra cosa distinta del azar puede estar detrás de ello. Este es precisamente el tipo de razonamiento que anima los esfuerzos del lunático por demostrar que la pauta de repetición de las vocales en las obras de Shakespeare se puede explicar mediante la hipótesis de que estaba escribiendo el nombre de la Bestia 666 veces, o cualquier otra cosa por el estilo. Para decirlo brevemente, el mero hecho de que un resultado sea improbable no debería animarnos a buscar algún tipo de explicación especial. Necesitamos alguna razón más para pensar que este resultado improbable puede deberse a algo más que el azar. El azar es capaz de producir tantas improbabilidades como el diseño.

Explicaciones y paradigmas

El proceso inductivo consiste en contemplar las cosas que pertenecen a nuestra experiencia como si fueran representativas del mundo que hay más allá de ella. Se trata de una proyección o extrapolación. Pero no es más que una parte de un proceso más amplio que pretende aumentar nuestro conocimiento de las cosas. En la última sección de este capítulo voy a presentar algunos de los razonamientos que forman parte de este proceso.

Supongamos que tenemos un sistema complejo. Contiene varios

elementos que parecen interactuar entre ellos. Podemos descubrir la forma en que lo hacen a través de los cambios y las variaciones que experimentan. Podríamos incluso ser capaces de entrelazarlos unos con otros y encontrar relaciones fiables. Un ejemplo de ello sería la ley de Boyle, según la cual la presión de una determinada masa de gas es inversamente proporcional a su volumen, dada una cierta temperatura. Se trata de una ley puramente empírica. Se ha comprobado que corresponde a la experiencia, y pensamos que se mantiene a lo ancho del mundo. Algunas disciplinas se sentirían terriblemente satisfechas si consiguieran llegar tan lejos. La economía, por ejemplo, pretende encontrar los elementos que realmente definen los sistemas económicos y ser capaz de describir las relaciones que se establecen entre ellos de un modo fiable, lo cual se revela muy difícil. Hace falta talento y esfuerzo, y la mayoría de intentos terminan en catástrofe. Tenemos la costumbre de olvidar que la física se encontraba en la misma situación. Por ejemplo, hizo falta un siglo de esfuerzos por parte de los científicos para descubrir que la energía es el factor fundamental de un sistema mecánico, ya que el hecho de que se conserve les permite prever el comportamiento del sistema. Se trata de un hecho histórico que los profesores de ciencia deberían escribir cien veces de forma obligatoria para que perdieran la costumbre de calificar de «estúpido» a un niño que no cae en ello de forma inmediata.

Si un economista presentara una propuesta acerca de cuáles son las variables correctas y las relaciones que se establecen entre ellas, podemos decir que tendríamos un modelo de la economía. Pero incluso si dispusiéramos de una propuesta de tales características, podría ser que continuáramos teniendo la impresión de que no comprendemos realmente lo que sucede. Isaac Newton (1642-1727) encontró una ley que describía la atracción gravitacional que existe entre los cuerpos en función de su masa y de la distancia que los separa: la famosa ley del cuadrado inverso. Pero tanto él como sus

contemporáneos tenían la impresión de que todavía les faltaba conseguir una comprensión real de por qué la gravedad actúa de este modo. Lo mismo sucede con la ley de Boyle: cuando tenemos una información de este tipo, podemos decir que sabemos algo acerca del sistema, pero no comprendemos realmente por qué se comporta de este modo. ¿Por qué habría de variar la presión en proporción inversa al volumen? ¿Por qué habría de ser siempre así? ¿Y por qué habría de tener alguna importancia que la temperatura sea constante?

Tales preguntas fueron contestadas por un modelo más sólido. La teoría cinética de los gases contempla los gases como volúmenes de moléculas en movimiento. La presión es el resultado del impacto de tales moléculas contra las paredes del recipiente. El movimiento de las moléculas se acelera cuando aumenta la temperatura. Cuando contemplamos los gases desde esta perspectiva, nos damos cuenta de que se trata de un mecanismo y de que, por lo tanto, si realizamos las suposiciones adecuadas, la propia naturaleza del mecanismo nos permitirá deducir las leyes empíricas que lo gobiernan, como por ejemplo la ley de Boyle.

Descubrir un mecanismo no significa evitar el problema de la inducción. El comportamiento continuo y uniforme de los elementos del mecanismo es una proyección o extrapolación realizada a partir de lo que hemos visto hasta el momento, igual que en cualquier otro caso. Sin embargo, reduce el número de suposiciones independientes que debemos realizar. Si las cosas poseyeran un conjunto de rasgos estables y una relación fiable entre ellos, tal vez los demás se podrían explicar a partir de ellos. Si damos por supuesta la estabilidad de tales rasgos, estamos en condiciones de explicar los otros en términos de éstos. En esto consisten los ideales de explicación y simplificación a los que aspira la ciencia.

Pero ¿cuándo podemos considerar que un «mecanismo» es satisfactorio? ¿Cuando comprendemos el comportamiento de sus elementos de forma «clara y distinta»? ¿O tal vez depende de otra

cosa? El intento de encontrar una respuesta para esta pregunta abre uno de los capítulos más excitantes del pensamiento moderno. Prácticamente todo el mundo se inclina a pensar que comprendemos mejor unos sistemas que otros. La mayor parte de la gente piensa que hay algunas relaciones causales que resultan especialmente inteligibles, como por ejemplo las colisiones entre objetos, mientras que hay otras, como las acciones que se producen a distancia o la influencia del cuerpo sobre la mente, que resultan altamente misteriosas. En realidad, esto era lo que pensaba casi todo el mundo —incluidos los filósofos y los científicos como Newton— antes de la llegada de Hume. Creían que poseíamos un conocimiento *a priori* acerca de cuál era la causa de tal cosa, y todavía más, de qué tipo de cosas no podían serlo. Se trata de algo que ya hemos visto antes. Incluso Newton consideraba evidente que la atracción gravitacional no podía consistir en una acción a distancia. Pensaba que cualquier idiota podía ver que si el Sol ejercía una atracción sobre la Tierra tenía que ser porque existía algún tipo de cadena causal que los relacionaba. La causalidad consistía necesariamente en tirones y empujones:

Que la gravedad pueda ser una propiedad connatural, inherente y esencial de la materia, de modo que un cuerpo pueda actuar a distancia sobre otro a través del vacío, sin ninguna otra mediación que haga posible la transmisión de la acción y la fuerza del uno al otro, me parece una absurdidad tan enorme que estoy convencido de que nunca se le ocurriría a nadie que tuviera un juicio competente en cuestiones científicas.⁴

Pensar que un cuerpo no puede actuar en un lugar distinto de aquel en el que se encuentra resulta seguramente «ajustado a la razón», o bien «claro y distinto» *o a priori*. Todavía razonamos de este modo cuando intentamos demostrar mediante argumentos puramente racionales que el universo tiene que ser obra de un dios,

por ejemplo. Seguimos creyendo que podemos saber qué tipo de cosas deben provocar ciertos efectos y qué tipo de cosas no.

Hume desmonta completamente esta clase de racionalismo:

Me permitiré afirmar, como proposición general que no admite excepción, que el conocimiento de esta relación en ningún caso se alcanza por razonamientos *a priori*, sino que surge enteramente de la experiencia, cuando encontramos que objetos particulares cualesquiera están constantemente unidos entre sí. Preséntese un objeto a un hombre muy bien dotado de razón y luces naturales. Si este objeto le fuera enteramente nuevo, no sería capaz, ni por el más meticuloso estudio de sus cualidades sensibles, de descubrir cualesquiera de sus causas o efectos. Adán, aun en el caso de que le concediésemos facultades racionales totalmente desarrolladas desde su nacimiento, no habría podido inferir de la fluidez y transparencia del agua, que le podría ahogar, o de la luz y el calor del fuego, que le podría consumir.⁵

Tal como haría un buen psicólogo, Hume ofrece también una explicación para el prejuicio de que podemos discurrir *a priori* acerca de las relaciones de causa-efecto:

Nos imaginamos que si de improviso nos encontráramos en este mundo, podríamos desde el primer momento inferir que una bola de billar comunica su moción a otra al impulsarla, y que no tendríamos que esperar el suceso para pronunciarnos con certeza acerca de él. Tal es el influjo del hábito que, donde es más fuerte, además de compensar nuestra ignorancia, incluso se oculta y parece no darse meramente porque se da en grado sumo.⁶

Hume sabía que los filósofos y los científicos aspiraban al ideal de conseguir algún tipo de «intuición» de las leyes de la naturaleza: algo parecido a una geometría o un álgebra que les permitiera comprender la sucesión de los hechos en función de unos modelos dotados de una certeza y una necesidad matemáticas. Aspiraban a una percepción «clara y distinta» de tipo cartesiano de por qué las

cosas son como son. Pero Hume cree que este objetivo es una ilusión. Haga lo que haga el científico, está más allá de su alcance.

Es bueno recordar en este punto que cuando Newton publicó en 1687 los *Principia Mathematica*, en los que describía las leyes del movimiento, hubo algunos científicos que se sintieron defraudados. Deseaban comprender en qué consistía la atracción gravitatoria, y Newton sólo les decía cómo funcionaba. Newton nos dice cómo acelera un cuerpo en relación con otro, y nada más que eso. Hume pensaba que la clase de teoría que había desarrollado Newton era la única que la ciencia podría conseguir jamás. Cualquier otra cosa que se pretendiera conseguir era, en opinión de Hume, un ideal inconsistente. En el pasaje que veremos a continuación, el término «filósofos» se refiere a los científicos, y «filosofía de corte natural» a lo que nosotros llamaríamos «ciencia natural», y en especial la física y la química:

Con esto podemos descubrir la razón por la que ningún filósofo, que sea razonable y modesto, ha intentado mostrar la causa última de cualquier operación natural o exponer con claridad la acción de la fuerza que produce cualquier efecto singular en el universo. Se reconoce que el mayor esfuerzo de la razón humana consiste en reducir los principios productivos de los fenómenos naturales a una mayor simplicidad, y los muchos efectos particulares a unos pocos generales por medio de razonamientos apoyados en la analogía, la experiencia y la observación. Pero, en lo que concierne a las causas de estas causas generales, vanamente intentaríamos su descubrimiento, ni podemos satisfacernos jamás con cualquier explicación de ellas. Estas fuentes y principios últimos están totalmente vedados a la curiosidad e investigación humanas. Elasticidad, gravedad, cohesión de partes y comunicación del movimiento mediante el impulso: éstas son probablemente las causas y principios últimos que podremos llegar a descubrir en la naturaleza. Y nos podemos considerar suficientemente afortunados, si somos capaces, mediante la investigación meticulosa y el razonamiento, de elevar los fenómenos naturales hasta estos principios generales, o aproximarlos a

ellos. La más perfecta filosofía de corte natural sólo despeja un poco nuestra ignorancia.⁷

Lo que tenemos aquí es una magnífica refutación del ideal racionalista. En su lugar, parece que lo único que nos queda es una mayor o menor familiaridad con ciertos sistemas. Cuando nos sentimos cómodos con un determinado sistema lo empleamos como «paradigma» o como un sistema que sirve como modelo de referencia para los demás. Nuestra idea de aquello que puede valer como una explicación satisfactoria se basa en estos paradigmas. Pero ahora que ya no podemos confiar en el ideal racionalista, nos damos cuenta de que esta misma idea puede ser mudable. Una vez que hemos sustituido la «razón» por el «hábito y costumbre», ¿acaso no pueden cambiar nuestros hábitos y costumbres? El famoso filósofo de la ciencia Thomas Kuhn (1922-1996) estaba convencido de que sí.⁸ La ciencia «normal» trabaja a partir de un conjunto de «paradigmas» o presuposiciones acerca de qué tipo de explicaciones debemos tratar de encontrar. La ciencia conoce períodos revolucionarios en los que se cuestionan los propios paradigmas. Debemos concebir la ciencia como «una sucesión de períodos de paz sacudidos por violentas revoluciones intelectuales». Después de cada revolución, cambia el tipo de explicaciones de las cosas que nos hacen sentir cómodos.

Algunas personas se excitan con bastante facilidad ante este tipo de planteamientos. Creen que sugieren algún tipo de «relativismo», según el cual cada uno tendría sus propios «paradigmas» y no habría forma de juzgar cuál de ellos es mejor. Pero tal sospecha resulta injustificada. Los paradigmas pueden ser mejores o peores. Contemplar el cielo estrellado como un velo opaco lleno de agujeros a través de los cuales nos llegan atisbos del cielo que se encuentra más allá de él fue durante una época el paradigma o el modelo dominante acerca de la naturaleza del cielo. Estamos convencidos de que el nuestro es mejor, y yo mismo sostengo también esta creencia. Es

perfectamente válido someter a juicio el valor de cada paradigma, y algunos de ellos ciertamente demuestran no tener ninguno.

Así pues, supongamos que fuera cierto que contemplamos inevitablemente el mundo a partir de un conjunto de categorías privilegiadas, en parte definidas por nuestra cultura e historia. Sigue sin poderse deducir de ello que todas las categorías sean igualmente «buenas». Algunos de estos conjuntos han sido descartados por razones adecuadas y suficientes. Un marco de investigaciones científicas consiste (idealmente) en un proceso constante de experimentación, predicción y verificación que tiende a descartar nuestros planteamientos erróneos. Sólo aquellos que logren sobrevivir a este proceso pasarán a la siguiente generación. Esto no significa que los marcos actuales de investigación científica correspondan a un modelo tan ideal: la ciencia todavía cae en todo tipo de omisiones, prejuicios y distorsiones. Pero el procedimiento científico contiene en sí mismo los mecanismos necesarios para corregirlos. Deberíamos recordar aquí el debate que planteamos en el primer capítulo, cuando criticábamos a Descartes por no tener en cuenta la naturaleza «autocorrectora» del conocimiento sensible, que nos permite reconocer las ilusiones como tales. Del mismo modo, la ciencia posee los mecanismos necesarios para corregir las ilusiones en las que ella misma cae. Toda su gloria procede de esto. Cuando nos encontramos ante construcciones especulativas que no contienen tales mecanismos —podríamos citar en este punto el psicoanálisis, las grandes teorías políticas, la ciencia de la «New Age» o la ciencia creacionista— podemos permitirnos el lujo de reservar nuestro interés para mejores ocasiones.

En el presente capítulo hemos descubierto algunos de los elementos que debemos tener en cuenta en nuestros razonamientos. Hemos tenido ocasión de ver algunas de las ideas en las que se basa la lógica formal. Hemos distinguido algunos procesos de razonamiento inductivo, y señalado su dependencia en relación con la

fe en la uniformidad de la naturaleza. Tenemos una idea acerca de cómo hay que razonar en relación con las diversas probabilidades de los sucesos. Hemos echado una ojeada a los procesos de construcción de modelos y explicaciones y hemos acabado por desconfiar de los razonamientos *a priori* acerca de las relaciones de causa-efecto. Todo ello nos dará herramientas para pensar acerca del mundo y del lugar que ocupamos en él.

CAPÍTULO

7

El mundo

En los seis capítulos que hemos visto hasta ahora, hemos tenido ocasión de explorar seis áreas de problemas. La lista es la siguiente: problemas relativos a nuestra fiabilidad en general; problemas relativos a la relación entre la mente y el cuerpo; problemas relativos a la libertad y el destino; problemas relativos al yo; problemas relativos a Dios; y problemas relativos al orden de la naturaleza. Cada una de estas áreas es notoria por las dificultades que plantea, y no resulta fácil encontrar la forma adecuada de pensar, o de estructurar los conceptos. Tal vez creamos que reflexionar acerca del mundo que nos rodea es una tarea relativamente exenta de problemas, en comparación con lo anterior. En el presente capítulo tendremos ocasión de tratar algunas áreas en las que un poco de reflexión también plantea numerosas dificultades.

Colores, olores, sonidos, tactos y sabores

Comenzaremos con una cita del gran físico Galileo Galilei (1564-1642), contemporáneo de Descartes:

Por lo tanto, digo, que tan pronto como concibo una materia o sustancia corpórea, experimento la necesidad de imaginar, al mismo tiempo, que está delimitada y que tiene esta o aquella forma, que con relación a otras es grande o pequeña, que está en este o aquel lugar, en este o aquel tiempo, que se mueve o está quieta, que toca o no toca a otro cuerpo, que es una, pocas o muchas, y ni con gran imaginación puedo separarla de estas condiciones. Pero que tenga que ser blanca o roja, amarga o dulce, sonora o muda, de olor agradable o desagradable, no siento que tenga que obligar a mi mente a representármela acompañada necesariamente de tales condiciones; más bien, si los sentidos no las hubieran advertido, tal vez la razón o la imaginación por sí mismas no lo hubieran logrado nunca. Por ello, pienso que estos sabores, olores, colores, etc., por parte del sujeto en el que parece que residen no son otra cosa que puros nombres y que tienen únicamente su residencia en el cuerpo sensitivo, de manera que eliminado el animal sensitivo, se eliminan y aniquilan todas estas cualidades.¹

Galileo está describiendo lo que más tarde se conocería como la distinción entre las cualidades primarias y secundarias de los objetos materiales. Las cualidades secundarias son percibidas de forma inmediata por los sentidos: son los colores, los sabores, los sonidos, los olores y los tactos. Según Galileo «residen» tan sólo en los animales sensibles (es decir, dotados de percepción). Y de acuerdo con Descartes, no hay razón para pensar que deban «parecerse» a su causa natural, sea cual sea (en el caso del color, por ejemplo, los fotones que entran en contacto con el ojo).

Descartes tenía un concepto igualmente pobre de los sentidos en cuanto vehículos de impresiones verdaderas (recordemos la bola de cera del capítulo 1):

[T]ales percepciones sólo me las ha dado la naturaleza para indicarle a la mente qué cosas son cómodas o incómodas para el compuesto del que ella forma parte [...] [pero] me he acostumbrado a alterar el orden de la naturaleza, porque uso las percepciones de los sentidos como si fueran reglas ciertas para discernir cuál es la esencia de los cuerpos

externos.²

Un ejemplo que le gustaba especialmente a Descartes era el de la percepción de un dolor en el pie, que se produce después de que los «espíritus animales» transporten al cerebro su energía (su «chorro de llamas»), a través de los nervios. Dios lo dispuso de manera que la mente recibiera la mejor sensación posible (de nuevo, el buen arbitrio divino). La actividad propia del cerebro podría haber transmitido algo distinto a la mente. Sin embargo, «ninguna otra cosa hubiera contribuido tanto al bienestar continuado del cuerpo». En otras palabras, si Dios hubiera dispuesto que yo interpretara la actividad del cerebro como si indicara simplemente una perturbación en el propio cerebro, por ejemplo, no apartaría con la suficiente rapidez el pie que está sufriendo el daño de la fuente del dolor. Démonos cuenta de que la postura de Descartes en este punto rechaza de forma implícita el epifenomenalismo. El hecho de que apartemos rápidamente el pie o no depende de que se produzcan unos u otros estados mentales. Si lo mental no tuviera ningún efecto, Dios podría haber dejado que reaccionara de cualquier forma sin que ello tuviera la menor influencia sobre nuestro bienestar.

Detrás del pintoresco lenguaje que emplea Descartes se esconde una idea sorprendentemente moderna. Si sustituimos a Dios por la evolución, podemos interpretar el pasaje de la siguiente forma: para que una criatura pueda prosperar es necesario que reciba de su entorno la información precisa para satisfacer sus necesidades actuales. Para ello debe recibir una información que la estimule a realizar las acciones adecuadas. Por ejemplo, en caso de que se acerque un depredador, necesita recibir algún tipo de información que la estimule a huir. A estos efectos, sin embargo, no importa cómo sea la sensación que experimente. Si el depredador pisa una rama caída en el suelo, puede que los «movimientos» que eso provoque en el oído hagan que la criatura oiga un ruido, sea fuerte, grave, agudo,

discordante o armónico, o bien que perciba simplemente un mal olor, pero mientras sienta algo que le infunda temor, los sentidos están cumpliendo con su cometido. Los sentidos sólo nos ofrecen datos «confusos» (en oposición a lo que podríamos considerar datos claros y distintos).

A su vez, esta circunstancia deja a Dios libre de toda culpa. Si nos preguntamos por qué nos engañan los sentidos al hacernos pensar que existen colores en el mundo exterior, la respuesta vendrá en dos partes. En primer lugar, Dios creó «el mejor sistema que se podía concebir» con el fin de que la sensación fuera la más indicada para preservar la salud de la persona. Tal como hemos visto, los sentidos pueden engañarnos. Pero Descartes insiste en que somos nosotros, y no el plan de Dios o el de la Naturaleza, quienes cometemos el error de malinterpretar los datos procedentes de los sentidos al no corregirlos mediante la inteligencia. No deberíamos considerar que los datos de los sentidos nos aportan una información inmediatamente correcta acerca de las propiedades reales de las cosas. Ello significaría tratar los datos de los sentidos como si fueran claros y distintos, en lugar de confusos.

Cuando empleamos el intelecto para abstraemos de los datos que nos ofrecen los sentidos, ¿con qué clase de mundo nos quedamos? El matemático Descartes creía que la auténtica propiedad de la *res extensa* era precisamente la extensión. Todo lo demás venía marcado por la actividad posiblemente engañosa de los sentidos, que «rellenan» los volúmenes espaciales con cosas como colores y tactos: cosas que Descartes, al igual que Galileo, pensaba que residían tan sólo en la mente. Así pues, del mismo modo que fueron Descartes y sus contemporáneos quienes plantearon el dualismo entre mente y cuerpo, también fueron ellos quienes plantearon el dualismo entre el mundo tal como es para nosotros —el mundo familiar, cálido y confortable, lleno de colores, olores y sonidos— y el mundo tal como es objetivamente, o en sentido absoluto (su «imagen científica»): un

mundo que sólo contiene partículas físicas y fuerzas que actúan a lo largo de los espacios ilimitados del cosmos.

¿Por qué pensamos que la ciencia reduce los colores y las demás sensaciones a una existencia puramente mental? El argumento más convincente en este sentido proviene de la relatividad de la percepción. La gente acostumbra a pensar en el «relativismo» como una amenaza o una tentación que tiene una presencia especial en el campo de la filosofía moral, ya que estamos familiarizados con las incómodas diferencias que se producen en las reacciones de distintas personas ante una misma situación. Pero en este punto se plantea un relativismo de corte más general. El argumento acerca de la relatividad de los sabores, los olores, los colores, los sonidos y los «tactos» se puede plantear del siguiente modo:

Supongamos que una parte del mundo o un objeto que forma parte de él tiene un determinado olor, etc., para un observador O. Dicho olor, etc., dependen de las particulares estructuras sensoriales de O. *Luego*, debemos aceptar que existe o puede existir otra criatura O' que posea unas estructuras sensoriales distintas, para la que esta misma parte del mundo o este objeto tendrá un olor, etc., distintos. O y O' pueden llevar vidas igualmente eficientes y adaptadas. *Luego*, no hay *razón* para pensar que tan sólo uno de los dos percibe correctamente los olores, etc. *Luego*, no hay una única distribución correcta de los olores, etc., en el mundo. *Luego*, es mejor concebir los olores, etc., como algo que depende por completo de la mente.

Este argumento ya era habitual en el mundo antiguo, antes de que recuperara otra vez su protagonismo en el siglo XVII. Es importante señalar algunas cuestiones acerca de él.

Primero, este hecho no depende de la existencia real de una criatura distinta (O'). Es suficiente con que podamos concebir la posibilidad de la existencia de tal criatura: alguien que tuviera una sensibilidad distinta al color, o cuyo aparato auditivo respondiera a distintas frecuencias de sonido, o más en general a energías de

cualquier otra clase. Por supuesto, el argumento resulta más convincente cuando encontramos ejemplos llamativos. Nadie que posea un perro puede creer realmente que el mundo de olores en el que él vive se parece en algo al que conocemos nosotros. Y muchos de nosotros recordamos perfectamente, por ejemplo, el distinto sabor que tenía la cerveza o el vino antes de que nos acostumbráramos a ellos. Es indudable que existen distintas sensibilidades: para empezar, la mayoría de los mamíferos son insensibles al color, con la excepción de algunos primates. Hay sustancias (el fenol thio-urea, por ejemplo) que tienen un sabor marcadamente amargo para la mayoría de los seres humanos y, sin embargo, no tienen ningún sabor para los demás. Y no nos faltarían ejemplos para seguir.

Sin embargo, con independencia de los ejemplos que se puedan señalar, resulta fácil imaginar otras formas de vida que pudieran sobrevivir sin problemas gracias a estructuras sensoriales significativamente distintas. Hay personas que ajustan sus televisores de modo que la imagen resulta excesivamente brillante y contrastada para los demás y, sin embargo, todos ven como resultado las mismas imágenes.

La segunda premisa también parece incontestable. Forma parte del conocimiento que todos tenemos acerca del mundo. Sabemos que bajo ciertas condiciones, por ejemplo, las cosas adquieren un sabor bien distinto. Cuando estamos resfriados perdemos buena parte de nuestro olfato. Tenemos un conocimiento bastante detallado del proceso que hace posible la visión, gracias a la distinta sensibilidad de los tres tipos de receptores que hay en el ojo, tal como vimos en el capítulo 2. Y sabemos que los murciélagos se orientan gracias a un sistema acústico que está vedado a los seres humanos.

Así pues, la conclusión inicial parece inevitable. Podemos compararla con la conclusión del siguiente argumento: dos equipos distintos de televisión reciben la misma señal. Sin embargo, la imagen que se pueda ver en cada uno de ellos depende de la

particular estructura de cada equipo. Por lo tanto, podría ser que distintos televisores ofrecieran resultados distintos a partir de la misma señal (y por supuesto eso es precisamente lo que sucede).

La siguiente premisa es crucial, aunque a menudo sea pasada por alto en algunas discusiones acerca del relativismo referido a otras áreas, como la ética. Cualquier argumento que pretenda llegar a algo parecido a la conclusión final tiene que pasar necesariamente por este punto. No sirve de nada señalar que existen distintas percepciones del mundo, si no cerramos el paso a la interpretación de que tan sólo una de ellas sea la correcta. La analogía con los televisores puede servir también para ilustrar este punto. Alguien podría decir con toda justicia que la imagen que muestra un determinado televisor en respuesta a la señal está sujeta a variaciones. Si el modelo del televisor no es el adecuado para el tipo de señal que recibe, por ejemplo, no veremos más que nieve. Pero la única conclusión que podemos sacar de ello es que el televisor no consigue interpretar una información que existe, que está objetivamente allí, en la señal. Y no vale como argumento decir que la información no existe con independencia del receptor. Desde el momento en que hay un transmisor que emite el mensaje original, un televisor que sólo muestre nieve cumple peor su función que otro que muestre el mensaje. No es que la cumpla bien, sino de un modo distinto. Y sin embargo esto es precisamente lo que pretende este argumento en favor del relativismo. Así pues, se trata de un argumento deficiente.

Y lo mismo sucedería con nuestro argumento si no fuera por la premisa de que las diferentes criaturas podrían llevar vidas igualmente funcionales y adaptadas. Así pues, para corregir aquel defecto debemos aceptar que la adaptación al mundo de los diferentes observadores sea potencialmente igual de buena. Para Descartes, esto último sería una creencia con un fundamento teológico. Para nosotros, la explicación podría ser de tipo

evolucionista. Aquellas criaturas que no reciben el tipo de información que necesitan para sobrevivir simplemente desaparecen. De modo que, a diferencia de lo que sucede en el caso de los televisores, O y O' pueden obtener unos resultados igual de buenos, aunque sus vidas se basen en experiencias sensoriales distintas: colores, olores, sonidos, sabores y tactos distintos. Y es precisamente esta igualdad la que sugiere que sería «favoritismo», para emplear la expresión de Russell, decir que una de las formas de representar el mundo es mejor que la otra.³

Sin embargo, tal vez no tengamos suficiente con la premisa de la adaptación equivalente. Podría ser que tuviéramos la tentación de pensar lo siguiente: los perros, por ejemplo, son unas criaturas ciertamente bien adaptadas, con unos sistemas sensoriales que se ajustan perfectamente a sus necesidades naturales. Pero consideremos las diferentes dimensiones de la experiencia sensorial. Los perros poseen un olfato maravilloso. Admitamos, pues, que son capaces de percibir olores que nosotros no podemos percibir. A ellos les corresponde toda la «autoridad» en materia de olores. Por otro lado, los perros son insensibles al color. Por lo tanto, no podemos considerar que tengan ninguna «autoridad» en materia de colores. Los seres humanos somos capaces de realizar discriminaciones visuales mucho más sutiles entre objetos, y bajo un espectro de luces mucho más completo que los perros. Para ello sirve precisamente nuestra visión de los colores. Así pues, ¿por qué no habríamos de decir que los colores reales son los que ven las criaturas mejor adaptadas a los colores? ¿Y los olores reales los que perciben las criaturas mejor adaptadas a los olores? Si realmente podemos decir esto, la siguiente conclusión de nuestro argumento no es válida.

Este hecho señala ciertamente una deficiencia en el argumento, tal como está planteado. Para corregirla se requeriría una premisa más fuerte. Un posible parche sería considerar cada una de las dimensiones sensoriales D (visión, tacto, olor, sonido, sabor) por

separado. En tal caso tendríamos cinco argumentos distintos, y en cada uno de ellos la premisa crucial tendría la siguiente forma:

O y O' pueden tener vidas igualmente eficientes y adaptadas en relación con la dimensión sensorial D.

Si aceptamos esto último, el resto del argumento viene a ser un paseo. El objetivo de la maniobra final es bien claro. El phenol thio-urea en sí no puede tener y no tener sabor a la vez. De modo parecido, no podemos pensar que el mundo contenga tantos olores como posibles aparatos sensoriales, en función de la diferente adaptación de éstos para captar ciertas moléculas (o su ausencia) en ciertas combinaciones y concentraciones. En un mundo así coexistirían un número infinito de olores, puesto que no hay límite para las variedades posibles de receptores. Como resultado de este argumento llegamos al llamado «idealismo de las cualidades secundarias». Significa llegar a la misma conclusión que Galileo, es decir, que los objetos inmediatos de la experiencia sensorial pertenecen «a la mente».

La solidez del sentido común

Esta conclusión no les pareció tan mala a Descartes y la mayoría de sus sucesores. Tal como hemos visto, Descartes contaba todavía con la «razón» para informarle de las propiedades que poseían realmente los objetos. No le preocupaba demasiado el aspecto ilusorio del mundo de la apariencia: el hecho de que los colores procedan, por así decirlo, de nosotros mismos antes que de los objetos que vemos. Esta postura fue consagrada dentro de la filosofía de habla inglesa por John Locke.

Locke es muy explícito en este punto. Existen

cualidades *originales* o *primarias* de un cuerpo, las cuales, creo,

podemos advertir que producen en nosotros las ideas simples de la solidez, la extensión, la forma, el movimiento, el reposo y el número.

Y también

Cualidades tales que en verdad no son nada en los objetos mismos, sino potencias para producir en nosotros diversas sensaciones por medio de sus cualidades primarias, es decir, por el bulto, la forma, la textura y el movimiento de sus partes insensibles, como son colores, sonidos, gustos, etc. A éstas llamo cualidades secundarias.⁴

Por un lado tenemos el mundo científico, integrado por lo que en la época de Locke se consideraba que eran los objetos reales: un mundo de pequeñas partículas que se acoplan unas a otras para formar cuerpos de mayor tamaño, cada uno de los cuales tendrá sus correspondientes cualidades primarias, es decir, científicas. Tal es la imagen científica del mundo. Por otro lado tenemos la imagen manifiesta: el mundo cálido o frío, lleno de colores, olores, sabores y ruidos en el que creemos vivir. Pero la imagen manifiesta reside tan sólo en la mente, o por lo menos se debe en gran medida a ella. No se puede decir lo mismo del mundo científico.

¿Acaso debemos entender que, según Locke, los objetos no tienen color u olor? En un sentido se puede decir que sí: los objetos tienen la capacidad de producir los colores y los sabores en nosotros. Sin embargo, poseer esta capacidad no significa tener tales colores u olores.

Cuanto he dicho tocante a los colores y olores, se puede entender también respecto a gustos, sonidos y demás cualidades sensibles semejantes, las cuales, cualesquiera que sea la realidad que equivocadamente les atribuimos, no son nada en verdad en los objetos mismos, sino potencias para producir en nosotros diversas sensaciones, y dependen de aquellas cualidades primarias, a saber: volumen, forma, textura y movimiento de sus partes, como ya dije.⁵

Las diversas sensaciones que experimentamos no guardan ningún

tipo de parecido con las causas que los producen.

El planteamiento de Locke es visto a menudo como un ejemplo de realismo puramente científico, natural y de sentido común. Podemos sustituir sus pequeñas partículas de materia, con sus formas y tamaños peculiares, por energías, fuerzas y campos o por partículas subatómicas. Sin embargo, la oposición esencial entre el mundo de la ciencia y su imagen aparente permanece en la mente de muchas personas tal como lo planteó Locke.

¿Acaso tiene eso algún problema?

Los problemas de Berkeley

Algunos escritores franceses no estaban del todo satisfechos con el planteamiento de Descartes. Después de todo, si Dios nos estaba engañando (aunque fuera por nuestro propio bien, por supuesto) en relación con las cualidades secundarias: ¿por qué no habría de hacerlo también en relación con las cualidades primarias? Si desde su punto de vista es bueno que veamos colores, aunque éstos no guarden la menor relación con la realidad física, ¿por qué no habría de serlo que viéramos objetos extensos, aunque la realidad física no fuera realmente espacial? Como un caballo de Troya, los colores vuelven a introducir el escepticismo general del Genio Maligno que Descartes había creído derrotar.

Así pues, tenemos un problema epistemológico. Sin embargo, descubriremos un nuevo motivo de inquietud si examinamos los presupuestos metafísicos del mundo científico. Intentemos pensar acerca de aquello que encontramos en el espacio. Descartes había eliminado todas las cualidades de la realidad física excepto una: la extensión. Pero todo el mundo consideraba que esto era insostenible. La «extensión» es una cualidad completamente abstracta. Un metro cúbico de espacio es una cosa; un metro cúbico de espacio ocupado por un cuerpo es algo muy distinto. Debemos concebir la realidad

física en términos de objetos que ocupan un lugar en el espacio, no sólo en términos de espacio en sí.

Pues bien, podemos pensar, ningún problema. Locke reconoce propiedades tales como la «solidez» o el «movimiento». Sin embargo, el movimiento no nos servirá de nada a menos que seamos capaces de identificar objetos que se mueven. De modo que concentrémonos en los objetos. Sabemos que un volumen en el espacio contiene un objeto gracias a la resistencia que éste ofrece. En esto consiste la diferencia entre un metro cúbico de espacio que contiene granito y un metro cúbico de vacío. Así pues, ¿en qué consiste la solidez? Locke lo define así:

La idea de la solidez la recibimos por nuestro tacto; y surge de la resistencia que advertimos en un cuerpo a que cualquier otro cuerpo ocupe el lugar que posee, hasta que cede. No hay ninguna otra idea que recibamos más constantemente por vía de sensación que la de solidez. Sea que estemos en movimiento, sea en reposo, cualquiera que sea la posición en que estemos, siempre sentimos algo debajo de nosotros, algo que nos sostiene y que impide sumirnos aún más. Y los cuerpos que manejamos a diario nos hacen percibir que mientras están en nuestras manos, por una fuerza invencible, impiden que se acerquen las partes de nuestras manos que los oprimen. Eso que de tal manera impide el acercamiento de dos cuerpos, cuando se mueven el uno hacia el otro, llamo la solidez. [...] [S]i alguien piensa que será mejor llamarle impenetrabilidad, tiene mi venia. Sólo que me ha parecido el término solidez más propio para expresar esa idea, no porque acarrea en sí algo más de positivo que el término impenetrabilidad, que es negativo, y quizá sea más una consecuencia de la solidez que no la solidez misma. Ésta, entre todas, parece ser la idea más íntimamente unida con lo corpóreo y esencial a éste; de tal suerte que no es de encontrarse o de ser imaginada en ningún otro lugar, sino sólo en la materia.⁶

La solidez de los objetos parece reducirse a su «capacidad» para excluir a otros objetos de la porción de espacio que ocupan. Pero

¿Acaso podemos darnos por satisfechos con una concepción del mundo que tan sólo reconoce distintas regiones del espacio con distintas propiedades? ¿Acaso no necesitamos algo más, algún tipo de sustancia que posea estas propiedades?

Por lo menos Locke admite que podemos conocer la solidez, de modo que tal vez su propuesta sea válida. Parece claro que reconocemos la solidez gracias al tacto. Locke pone bastante énfasis en este punto:

Si alguien me pregunta ¿qué es la solidez? Lo remito a sus propios sentidos para que lo informen: tome entre sus manos un pedernal o un balón y trate de juntarlos, y así sabrá. Si no le parece ésta una explicación suficiente de la solidez, qué cosa sea, y en qué consiste, yo le prometo decirle qué cosa es y en qué consiste, cuando él me diga qué es pensar y en qué consiste, o cuando me explique qué es la extensión o el movimiento, lo cual, quizá, parece más fácil.⁷

El problema que se esconde detrás de este planteamiento lanzaba ya sus primeros avisos en la época de Locke, aunque éste no fuera demasiado sensible a ellos, y estalló finalmente a comienzos del siglo XVIII en los escritos del francés Pierre Bayle (1647-1706) y en los del filósofo irlandés George Berkeley (1685-1753)- Berkeley plantea una serie de objeciones devastadoras frente a las ideas de Descartes y Locke. Su postura es que estas ideas no son coherentes desde el punto de vista metafísico ni desde el epistemológico. Su crítica es compleja y funciona a muchos niveles distintos, pero podremos apreciar su alcance general si la resumimos en dos apartados.

1. Examinemos de nuevo la concepción de Locke acerca de nuestra forma de reconocer la solidez. Si esto es todo cuanto podemos decir, ¿por qué habríamos de considerar que la solidez se encuentra en mejor posición que el color, el tacto o el calor? Si tales sensaciones no nos proporcionan ninguna idea válida acerca de las

cualidades reales de las cosas, y no son más que el resultado de ciertas propiedades de los objetos reales, ¿por qué habría de ser distinto en el caso de la solidez? ¿Cómo podemos pasar de la sensación de solidez que recibe la mente a una propiedad parecida que exista realmente en el mundo? Sea lo que sea la solidez «para la mente», no tiene por qué ser la misma solidez que existe en el mundo. Si nuestras ideas no son sólidas, ¿qué sentido tiene decir que se «parecen» a objetos sólidos?

Y si la solidez desaparece del mundo real, ¿qué nos queda? La respuesta de Berkeley se hizo célebre: nada. El mundo queda reducido por completo a la mente. Esta doctrina que se conoce como «idealismo subjetivo».

2. El planteamiento de Locke parece requerir una concepción del mundo en términos de puras cualidades primarias, limpio de todo aquello que según él reside en la mente. Pero ¿es posible? La posición de Berkeley es clara: «Niego que pueda abstraer unas de otras, o concebir de forma separada, unas cualidades que es imposible que existan con tal independencia». Pensemos en un objeto físico corriente, como por ejemplo un tomate. «Hagamos abstracción» de su color, tacto, olor, sabor y de las sensaciones que experimentamos al tenerlo entre las manos. ¿Qué nos queda? Un tomate perfectamente invisible, intangible, imposible de detectar: seguramente algo muy cercano a nada. Hume expresó la objeción en magníficos términos (en la siguiente cita, «la filosofía moderna» corresponde a la concepción de Locke):

La idea de solidez es la de dos objetos que, aun impelidos por la fuerza más extrema, no pueden, interpenetrarse, sino que siguen conservando una existencia separada y distinta. Por consiguiente, la solidez es perfectamente incomprensible de forma aislada y si no concebimos algunos cuerpos sólidos y manteniendo esta existencia separada y distinta. ¿Y cuál es la idea que tenemos de estos cuerpos? Adviértase que las ideas de colores, sonidos y otras cualidades segundas

han sido excluidas. Además, la idea de movimiento depende de la de extensión, y ésta de la de solidez; luego es imposible que la idea de solidez pueda depender de cualquiera de ellas, porque eso sería un círculo vicioso, haciendo que una idea dependa de otra, y que a la vez ésta dependa de aquélla. Nuestra moderna filosofía, pues, no nos proporciona ninguna idea, ni correcta ni convincente, de la solidez; por consiguiente, tampoco nos proporciona idea alguna de la materia.⁸

O en otras palabras: «[D]espués de excluir colores, sonidos, calor y frío del rango de existencias externas no queda nada que pueda proporcionarnos una idea precisa y consistente de los cuerpos».⁹ Tanto Berkeley como Hume niegan que podamos comprender las supuestas propiedades del supuesto mundo exterior, a menos que sea en términos de nuestra propia experiencia, es decir, de nuestra mente. La «filosofía moderna» o la concepción científica del mundo requiere que le demos sentido a una idea «absoluta» o «científica» de la realidad. De acuerdo con ella, la realidad es concebida en términos de objetos independientes que ocupan un lugar en el espacio, cuya estructura nos dice todo cuanto se puede decir acerca del universo entero, incluidos nosotros y nuestra experiencia. Pero si resulta que esta concepción contiene un error de base, tendremos que buscar alguna otra.

Fuerzas, campos y objetos

En el pasaje citado, Hume preguntaba qué concepción nos quedaba de aquellos objetos impenetrables, y argumentaba que sin el «relleno» aportado por las cualidades secundarias, no había respuesta posible, lo cual plantea un problema de tipo más general: ¿acaso tenemos alguna concepción de los cuerpos más allá de las propiedades que manifiestan en sus interacciones entre ellos y con nosotros?

No es éste el lugar adecuado para entrar en los detalles de la

física, pero tal vez podremos ganar algo de perspectiva a través de las palabras de una de las figuras más destacadas de esta ciencia, Michael Faraday (1791-1867). Supongamos que estamos intentando distinguir una partícula física a de las fuerzas o poderes m que nos permiten reconocer su influencia. Faraday escribe:

Tal como lo entiendo yo, [...] el núcleo o a desaparece, y la sustancia pasa a consistir en sus poderes, o m ; pues, ¿qué idea podemos formarnos acerca de este núcleo con independencia de sus poderes?: ¿En qué podemos basar nuestra suposición de un a independiente de las fuerzas que podemos reconocer? ¿Por qué debemos asumir entonces la existencia de algo sobre lo que somos perfectamente ignorantes, algo que somos incapaces de concebir y que no es en absoluto necesario desde el punto de vista filosófico?¹⁰

Faraday toma por los cuernos, por así decirlo, la anterior objeción de Hume acerca de la naturaleza de los cuerpos más allá de la solidez. Simplemente, no tenemos por qué pensar en las cosas como si fueran algo distinto de sus propiedades.

En el mundo de la física, la «imagen científica» se resuelve en un vasto flujo de fuerzas, que según las concebimos hoy serían las fuerzas gravitatorias, las atracciones y repulsiones electromagnéticas, y si vamos todavía un poco más allá, las interacciones fuertes y débiles entre partículas elementales. Pero llegados a este punto la situación se vuelve algo incómoda: recordemos que el argumento se plantea con la mayor de las generalidades, de modo que estas «partículas» se traducirán de nuevo en otras «fuerzas». El problema es que habitualmente, cuando pensamos en fuerzas o en cosas como campos magnéticos o gravitatorios, damos más o menos por supuesta la noción de una partícula. Comprendemos la existencia de un campo o de una fuerza en un punto del espacio en base a la aceleración que se produciría si introduyéramos en él una partícula experimental. Cuando tenemos un imán sobre la mesa, si disponemos otras «cosas» (partículas) a cierta distancia de él, vemos que

muestran una tendencia a moverse, y eso nos permite reconocer la existencia del campo magnético. En los típicos experimentos escolares, las partículas experimentales son limaduras de hierro.

Así pues, no podemos darnos por satisfechos con la imagen que se deriva de la propuesta de Faraday, según la cual las propias partículas se traducen de nuevo en otros poderes o fuerzas. Cuando se nos exige que intentemos comprender el cosmos sin la muleta que representan las «cosas», en el sentido más amplio, nos persigue de nuevo la objeción de Hume acerca de la impenetrabilidad: necesitamos saber en qué consiste eso que no se puede penetrar. Es como si la idea de sentido común de que existe una diferencia entre el espacio ocupado por un cuerpo y el espacio no ocupado hubiera sido abandonada por la idea de un espacio en el que se verifican ciertas relaciones condicionales (si...), en oposición a otros espacios en los que se verifican otras. Pero nosotros aspiramos a algo que realmente ocupe el espacio, y cuya presencia sea capaz de explicar la diferencia que existe entre las diversas relaciones condicionales, la diferencia que existe entre los diversos potenciales y poderes.

El problema se puede plantear igual que en el capítulo 2. Supongamos que Dios se dispone a crear el universo: ¿Cuánto trabajo tiene que hacer? ¿Puede crear sólo fuerzas? En tal caso, el universo parece convertirse en un gigantesco entramado de relaciones condicionales (si...). ¿O acaso también deberá crear objetos, por un lado para que las fuerzas tengan algo sobre lo que actuar y, por otro, quizá para que expliquen el origen de estas fuerzas? Si nos inclinamos por esto último, ¿qué concepción podemos tener de tales objetos? La primera posibilidad parece convertir el universo en un conjunto de potencialidades, como si fuera un inmenso juego de destellos. Tal vez Descartes, el matemático, se sentía atraído por esta imagen (la pregunta de si Descartes anticipó o no el planteamiento de Faraday resulta fascinante). Pero el sentido común parece exigir algo (algo sólido) que llene las porciones de espacio que contienen

materia.

Kant, uno de los primeros que propuso concebir la materia como un conjunto de «fuerzas», dedicó una atención especial a este problema. Según él, ésta era la mejor concepción de las cosas que podíamos alcanzar. En parte, el motivo es que percibimos el mundo a través de los sentidos, y éstos son esencialmente receptivos, es decir, no pueden ofrecernos más que el reflejo de ciertos poderes y fuerzas. Los sentidos no se han adaptado para decirnos qué hay detrás de la distribución de poderes y fuerzas en el espacio. Lo único que nos dan es el resultado de esta distribución. Cualquier cosa que hubiera detrás de ella tendría que ser enteramente «nouménica», es decir, tendría que estar más allá del alcance de las investigaciones científicas y, por eso mismo, más allá del alcance de la experiencia y el pensamiento humanos.

Hume pensaba que el problema de la impenetrabilidad ponía en entredicho toda la metafísica de «la filosofía moderna», aunque también pensaba que nadie podía tomarse en serio la retirada de Berkeley en el idealismo subjetivo. Kant también pensaba que el problema exigía una completa reformulación de la filosofía moderna.

Leyes y camisas de fuerza

La importancia de los problemas planteados en la última sección se descubre también si pensamos en otro elemento fundamental de la ciencia: el concepto de «ley de la naturaleza». Para ello debemos rescatar un ejemplo que vimos en el capítulo anterior: la lotería del arpa dorada.

Después de considerar el experimento mental, podríamos pensar algo parecido a lo siguiente: por muy impresionante que sea el experimento, tal vez nos esté llevando a cometer un error. El motivo es que representa las cosas como si el estado del mundo en cualquier período determinado fuera independiente de su estado en cualquier

otro. Es como si Dios lanzara un dado de seis caras al final de cada período, de modo que las opciones de que a continuación apareciera uno u otro color fueran de 1 entre 6. Si la situación fuera realmente ésa, sería indudablemente una falacia considerar que la repetición de un mismo número (correspondiente al azul) en cinco ocasiones convierte en más probable que se repita otra vez. Razonar de este modo significa caer en lo que se ha llamado «la falacia del jugador». No obstante, no parece que en el mundo tal como lo conocemos tenga lugar ninguna lotería de este estilo. No se produce el caos que cabría esperar si fuera así. Lo único que descubrimos son regularidades, de modo que resulta mucho más probable suponer que existe algo que garantiza un orden a través del tiempo. Nadie tira arbitrariamente un dado entre un momento y el siguiente: parece más bien como si Dios hubiera tomado una decisión y la mantuviera con firmeza. Tiene que haber una solución metafísica al problema de la inducción, aunque no se plantee en términos puramente matemáticos o probabilísticos.

Parece que esta observación puede sernos útil, pero ¿lo es realmente?

Por supuesto, parte del problema consiste en que incluso si el universo obedeciera a una única ley, por ejemplo una decisión de Dios, ésta podría ser perfectamente: «Sigamos el modelo T», en lugar de «Sigamos el modelo R». La ley invariable podría ser del tipo Torcido. Después de todo, nuestro conocimiento se limita a los segmentos de tiempo que han transcurrido hasta ahora. Y argumentar que puesto que la naturaleza ha mantenido una determinada uniformidad hasta ahora, es probable que siga manteniendo esta misma uniformidad, significa realizar una vez más, tal como señaló Hume, una inferencia inductiva.

Pero una vez más podemos descubrir un aspecto metafísico en el problema. Llamemos Camisa de Fuerza a aquello que garantiza un orden en la naturaleza. Una Camisa de Fuerza sería algo parecido a

una ley de la naturaleza que opera a través del tiempo: una directiva o una garantía que establece el orden de las cosas. La idea sería entonces que las cosas siguen funcionando igual a como lo han hecho toda la vida, por así decirlo, gracias a esta directiva o garantía. Y ahora el problema sería: ¿en qué podría consistir esta Camisa de Fuerza, si es que somos capaces de concebirla?

El problema que tenemos ahora resulta extremadamente parecido al problema que teníamos con el argumento cosmológico en el capítulo 5, e incluso podemos considerar que no es más que una versión de él. Las cosas que nos encontramos en el espacio y el tiempo, incluidas cosas tales como las decisiones humanas, son inherentemente mudables. Es posible que duren mucho tiempo, pero en la práctica van y vienen. Una Camisa de Fuerza no puede ser así, ya que si fuera mudable necesitaríamos alguna otra cosa que explicara su propia supervivencia a lo largo del tiempo, y de este modo nos veríamos arrastrados a una regresión.

Nuestra estrategia actual consiste en buscar el fundamento de las regularidades en alguna «otra cosa» que sea capaz de garantizar que la sucesión de los hechos seguirá siendo la misma. Pero entonces tenemos que dar razón de la continuidad y regularidad de esta «otra cosa». Si no es más que un «hecho dado», no ofrece mayores garantías que aquello que teníamos al principio: el orden empírico. Si necesita a su vez un fundamento distinto, nos vemos arrastrados a una regresión. Si decimos que es «necesario» o que tiene en sí mismo su propia explicación, nos enfrentamos al mismo escepticismo al que se enfrentaba el argumento cosmológico. No comprendemos realmente lo que significa una Camisa de Fuerza, ni tenemos criterios para señalar qué tipo de objetos corresponden a su definición.

En otras palabras, si una Camisa de Fuerza fuera algo tan mudable como todo lo demás, no tendríamos ningún motivo para confiar en su continuidad. Pero ¿podemos concebir algo que no esté sujeto al tiempo y al cambio? ¿Acaso podemos concebir algo siquiera parecido

a eso, ya no digo comprenderlo plenamente? ¿Acaso no estamos abocados una vez más a la amarga conclusión de Wittgenstein: «Una nada presta el mismo servicio que un algo sobre el que nada puede decirse»? O, en palabras de Hume:

El escenario del universo está continuamente cambiando y un objeto sigue a otro en sucesión ininterrumpida, pero se nos oculta absolutamente el poder o fuerza que mueve toda la máquina y jamás se revela ninguna de las cualidades sensibles del cuerpo.¹¹

Parece que nuestro entendimiento se ha quedado atascado una vez más. No somos capaces de concebir en qué consiste que una ley de la naturaleza se mantenga. Comprendemos el modo en que suceden las cosas, pero no tenemos la menor idea de por qué deben suceder de este modo.

En la sección anterior, siguiendo los pasos de Faraday y de Hume, vimos cómo la concepción científica «absoluta» de una realidad independiente tenía problemas para explicar la relación entre las cosas y sus poderes. Ahora descubrimos que nuestra concepción de estos mismos poderes que nos garantizan las leyes de la naturaleza es todo lo frágil que podía ser.

La revolución kantiana

Los problemas de la «filosofía moderna» llevaron a Berkeley a refugiarse en su propia mente. Decidió que el universo de nuestro entendimiento se limitaba a nuestras propias ideas y a nuestra propia naturaleza como «espíritus» o almas. Afortunadamente, no estamos solos en este universo subjetivo, ya que podemos estar seguros (pensaba Berkeley) de que nuestras experiencias proceden de un espíritu más poderoso: Dios (a estas alturas podemos imaginarnos perfectamente el guiño burlón de Hume ante este razonamiento causal). Sin embargo, nadie ha llegado a defender jamás que la

solución de Berkeley sea satisfactoria: da la impresión de que el Dios de Berkeley nos pone en la misma situación que el Genio Maligno de Descartes, es decir, nos enfrenta a una realidad virtual enteramente engañosa.

No obstante, hubo algunos filósofos que estuvieron de acuerdo con el diagnóstico inicial de Berkeley, y uno de ellos fue Kant. La idea de Kant era que la «filosofía moderna» de Locke había derivado hacia una postura insostenible, lo que él llamaba «realismo trascendental». Sería un «realismo» porque Locke insiste en la existencia de un mundo real de objetos independientes ubicados en el espacio y en el tiempo; y «trascendental» porque este mundo sería exterior a nuestra experiencia, y sólo podríamos conocerlo por medio de inferencias. Pero tanto Kant como Berkeley consideran que las inferencias son una base demasiado precaria. En el planteamiento de Locke:

No puedo, pues, hablando con propiedad, percibir Las cosas exteriores, sino sólo inferir su existencia desde mi percepción interna, considerando ésta como efecto de algo exterior que es su causa próxima. Ahora bien, el inferir una causa determinada a partir de un efecto dado es siempre inseguro, ya que el efecto puede proceder de más de una causa. En la relación de la percepción con su causa queda, pues, siempre la duda de si esta causa es interna o externa y, por consiguiente, de si todas las llamadas percepciones externas no son un simple juego de nuestro sentido interno; de si se refieren a verdaderos objetos exteriores como causas de las mismas. La existencia de tales objetos es simplemente inferida y corre el peligro de todas las deducciones. El objeto del sentido interno (yo mismo, con todas mis representaciones) es, en cambio, percibido inmediatamente y su existencia no ofrece dudas.¹²

Para Kant, lo más importante es escapar a este modelo del «teatro interior». Ya tuvimos ocasión de realizar una primera aproximación a su planteamiento en el capítulo 4, a propósito del yo. Entonces vimos que la autoconciencia es el resultado de una serie de hazañas

organizativas de estructura bastante compleja. Debemos organizar nuestra experiencia en términos de cierto orden temporal y espacial, para no caer en lo que Kant llama una mera «rapsodia» o caleidoscopio de percepciones.¹³ Sólo así podremos concebirnos a nosotros mismos como un sujeto que se mueve en un mundo de objetos independientes que ocupan un determinado espacio. ¿Qué intuición aporta Kant para superar el punto muerto al que había ido a parar la tradición filosófica posterior a Descartes?

Parte del mérito de Kant consiste en haber visto que la concepción del conocimiento que acabó defendiendo Locke era insostenible. Para Locke, el paradigma del conocimiento sería cierto «parecido» entre algo que se encuentra en la mente y los caracteres de los objetos que han sido su causa, como el que existe entre un objeto pintado en un cuadro y su modelo. Berkeley compartía este ideal. Es cierto que pensaba que tal parecido no podía cumplirse enteramente («Una idea no puede parecerse a otra cosa que no sea otra idea»). Sin embargo, su conclusión era que nuestro conocimiento se halla limitado al mundo de nuestras ideas. Kant se da cuenta de que un concepto no consiste en un cuadro mental, al menos en lo que se refiere al espacio y al tiempo, al tamaño, a la forma y al orden objetivo de las cosas. Consiste más bien en un principio organizativo o una regla; una forma de articular el flujo de información. El hecho de que todos usemos las mismas reglas o principios organizativos explica que todos tengamos la misma comprensión del mundo, a pesar de las diferencias que existen entre las diversas experiencias subjetivas.

Así pues, la conclusión sería que los problemas que hemos encontrado en las últimas secciones se deben a que buscábamos «cosas» que cumplieran determinados papeles: el papel de objetos que poseen ciertas fuerzas y poderes aunque sean independientes de ellos, o el papel de aquello que sustenta la causalidad y las leyes de la física. Pero si fuéramos capaces de liberar nuestra forma de pensar de esta dependencia de las cosas tal vez podríamos llegar más lejos.

Supongamos que concebimos nuestras ideas acerca de la causalidad y la legalidad, acerca de los objetos que se hallan en el espacio y en el tiempo, y acerca del espacio y el tiempo en sí mismos, como categorías necesarias del pensamiento. El entramado de principios que nos permiten organizar o sistematizar nuestra experiencia vendría dado por estas categorías. Ya no serían un conjunto de cosas que «inferimos» a partir de la experiencia. La idea que encontramos aquí es parecida a las ideas acerca del «yo» que ya tomamos prestadas de Kant, y podríamos decir que es la otra cara de la misma moneda. Si intentamos comprender el yo en términos sensibles, como si fuera un objeto más de la experiencia, nos enfrentamos a la objeción de Hume de que no hay tal objeto. En cambio, si pensamos que se trata de un punto de vista personal o egocéntrico que permite organizar la experiencia, el papel del yo en nuestro pensamiento se vuelve más claro, y lo mismo sucede con las ilusiones que había generado la idea del yo.

La revolución kantiana queda resumida en un famoso pasaje al comienzo de la *Crítica de la razón pura*:

Se ha supuesto hasta ahora que todo nuestro conocer debe regirse por los objetos. Sin embargo, todos los intentos realizados bajo tal supuesto con vistas a establecer *a priori*, mediante conceptos, algo sobre dichos objetos —algo que ampliara nuestro conocimiento— desembocaban en el fracaso. Intentemos, pues, por una vez, si no adelantaremos más en las tareas de la metafísica suponiendo que los objetos deben conformarse a nuestro conocimiento, cosa que concuerda ya mejor con la deseada posibilidad de un conocimiento *a priori* de dichos objetos, un conocimiento que pretende establecer algo sobre éstos antes de que nos sean dados. Ocurre aquí como con los primeros pensamientos de Copérnico. Éste, viendo que no conseguía explicar los movimientos celestes si aceptaba que todo el ejército de estrellas giraba alrededor del espectador, probó si no obtendría mejores resultados haciendo girar al espectador y dejando las estrellas en reposo. En la

metafísica se puede hacer el mismo ensayo, en lo que atañe a la intuición de los objetos. Si la intuición tuviera que regirse por la naturaleza de los objetos, no veo cómo podría conocerse algo *a priori* sobre esa naturaleza. Si, en cambio, es el objeto (en cuanto objeto de los sentidos) el que se rige por la naturaleza de nuestra facultad de intuición, puedo representarme fácilmente tal posibilidad.¹⁴

Esto es lo que Kant llama «idealismo trascendental». Kant pone un cuidado especial en distinguirlo del «idealismo subjetivo» de Berkeley. Y evidentemente tampoco puede ser el «realismo trascendental» de Locke. Luego ¿qué es?

Parece como si en el propio acto de tener una experiencia «creáramos» un mundo que debe conformarse a ella. La idea resulta muy extraña. Fue el universo el que nos creó a nosotros, hace algo así como 13.000 millones de años, y no nosotros al universo. Pero Kant no está intentando negar eso. Lo que quiere es comprender el modo en que ciertos conceptos, como el de cosa, fuerza, espacio, tiempo o causalidad, determinan la forma en que pensamos (y debemos pensar) acerca del mundo. Su intención no es negar ciertos elementos del pensamiento científico, o más concretamente del sentido común, sino explicar cómo se relacionan estos elementos en nuestro pensamiento. Son precisamente estos elementos los que estructuran lo que Kant llama el «mundo fenoménico»: el mundo que describe la ciencia y que se manifiesta en nuestra experiencia sensible.

Por supuesto, Kant no pensaba que nuestra experiencia sensible «creara» de algún modo este mundo. No lo pensaba, por ejemplo, en relación con las cualidades secundarias:

Los colores no son propiedades de los cuerpos a cuya intuición van ligados, sino que son simples modificaciones del sentido visual al ser éste afectado de alguna forma por la luz. El espacio, en cambio, pertenece necesariamente, en cuanto condición de los objetos externos,

al fenómeno o intuición de éstos. Ni el sabor ni los colores son condiciones indispensables para que puedan convertirse en objeto de nuestros sentidos.¹⁵

La idea es que el espacio, a diferencia del color, es una «condición» imprescindible para que algo pueda ser objeto de nuestra experiencia. La objetividad depende más del espacio que del color.

La principal dificultad a la hora de interpretar a Kant en este punto consiste en determinar si supera tanto las ideas de Berkeley como podría parecer. Supongamos que Berkeley le agradece a Kant tres intuiciones:

1. Debemos abandonar completamente la concepción del conocimiento en términos sensibles que defendía Locke, y contemplar los conceptos que empleamos para describir el mundo como reglas, principios y estructuras organizativas antes que como imágenes mentales.
2. Nuestra experiencia tiene que poseer un orden (de acuerdo con la expresión de un filósofo contemporáneo, Jonathan Bennet, tiene que haber un «límite de velocidad») para que nuestra autoconciencia sea posible.
3. Para que este orden sea posible, debemos ser conscientes de que ocupamos un determinado lugar en el espacio, desde el cual percibimos objetos permanentes en el espacio y en el tiempo, cuyo comportamiento se ajusta a ciertas pautas determinadas por las leyes de la naturaleza.

Podría parecer que todo esto no hace más que abonar el terreno para Berkeley. Éste también era consciente de que interpretamos nuestra experiencia en términos espacio-temporales, objetivos. Pero estaba convencido de que debíamos «hablar en vulgar por mucho que pensáramos con lo aprendido»: en otras palabras, debemos aprender a considerar nuestra forma de interpretar las cosas como

una cierta manera de hablar, más que como la descripción de un mundo real, independiente y objetivo.

Un factor que contribuye a aumentar todavía más la confusión es que Kant realiza afirmaciones que tienen una considerable afinidad con una postura no muy distinta a la del idealismo subjetivo de Berkeley. Su «revolución copernicana» le lleva a decir cosas como la siguiente:

En nuestro sistema, por el contrario, estas cosas exteriores, a saber, la materia en todas sus formas y modificaciones, no son más que fenómenos, es decir, representaciones nuestras de cuya realidad poseemos una conciencia inmediata.¹⁶

El elemento que aleja a Kant del idealismo subjetivo es su convicción de que tanto Descartes como sus seguidores habían tomado el rábano por las hojas. Ellos pensaban que la «experiencia interior» se mantenía sólida como una roca, mientras que el mundo exterior se volvía problemático. Para ir más allá de eso,

la prueba requerida debe, pues, mostrar que tenemos experiencia de las cosas externas, no simple imaginación. Ello no podrá ocurrir más que en el caso de que podamos demostrar que nuestra misma experiencia interna —indudable para Descartes— sólo es posible si suponemos la experiencia externa.¹⁷

«Experiencia externa» significa en este caso aquella experiencia que nos da una conciencia inmediata de una realidad que se extiende más allá de nosotros mismos. La cuestión de si Kant consigue resolver con éxito este problema, y cómo lo consigue, plantea uno de los debates más importantes del pensamiento moderno.

El ojo del espectador

Un auténtico realista, es decir, alguien que se opone al idealismo, pretende defender la existencia de hechos y estados de cosas

enteramente independientes de la mente. El idealista nos recuerda a cada momento el papel que juega la mente en la selección y elaboración de nuestra concepción del mundo. Para el idealista, la mente crea el mundo en el que vivimos, el *Lebenswelt** de nuestros pensamientos, imaginaciones y percepciones. Por supuesto, Kant se halla metido en esto hasta los codos, ya que todo el entramado desde el que pensamos, nuestro «esquema conceptual» del espacio, el tiempo, los objetos, las causas y los yoes, se debe a los principios organizativos de la mente.

Ahora bien, no hace falta ser un dualista cartesiano para simpatizar con este reconocimiento del trabajo que realiza la mente para generar el único mundo que somos capaces de comprender. En realidad, la mayoría de pensadores del siglo XX (siguiendo una tendencia del siglo XIX) ha recogido el testigo de Kant y ha corrido si cabe con mayor entusiasmo. En particular, lo que más ha celebrado ha sido aquello que antes he presentado con el nombre de «paradigmas»: la idea de que vemos el mundo a través de unas lentes que varían en función de la cultura y la historia, o bien que vivimos en palacios o prisiones conceptuales contruidos por nosotros mismos.

Una vez más, sin embargo, me gustaría introducir a los pensadores modernos a través de uno clásico, y de nuevo podríamos tomar a Berkeley como punto de partida. En el primero de sus *Tres diálogos* encontramos este célebre pasaje, en el que Filonús representa al propio Berkeley:

FILONÚS: [...] Pero (pasando por alto todo lo que se ha dicho hasta ahora y no teniéndolo en cuenta si así lo quieres), estoy dispuesto a jugarlo todo a una carta. Si puedes concebir que es posible que una mezcla o combinación de cualidades, o un objeto sensible cualquiera, existan fuera de la mente, admitiré que es así realmente.

HILAS: Si es así, pronto decidiremos la cuestión. No hay nada más fácil

* *Lebenswelt* significa «mundo de la vida» en alemán. [N. del t.]

que concebir un árbol o una casa que existen por sí mismos, independientes de cualquier mente y sin que los perciba mente alguna. En este momento los concibo perfectamente como existentes en esta forma.

FILONÚS: ¿Cómo dices, Hilas? ¿Puedes ver una cosa que al mismo tiempo no es vista?

HILAS: No, sería una contradicción.

FILONÚS: ¿No es una gran contradicción hablar de concebir una cosa que es inconcebible?

HILAS: Sí.

FILONÚS: El árbol o la casa de que hablas, son concebidos por ti.

HILAS: ¿Cómo podría ser de otro modo?

FILONÚS: Y, desde luego, lo que es concebido está en la mente.

HILAS: No hay duda de que lo que es concebido está en la mente.

FILONÚS: ¿Cómo puedes llegar a decir que concibes un árbol o una casa con existencia independiente de todas las mentes y exterior a ellas?

HILAS: Reconozco que ha habido un error; pero espera, déjame ver qué es lo que me ha inducido a error. Es curiosa la equivocación sufrida. Cuando estaba pensando en un árbol en un lugar solitario donde nadie estaba presente para verlo, me parecía que era concebir un árbol que existía sin ser percibido ni pensado, y no tenía en cuenta el hecho de que yo mismo lo concebía durante ese rato. Pero ahora veo claramente que todo lo que yo puedo hacer es forjarme ideas en mi propia mente. Puedo, sin duda, concebir en mis propios pensamientos la idea de un árbol, de una casa o de una montaña, pero eso es todo. Y esto está lejos de probar que puedo concebirlos existentes fuera de las mentes de todos los espíritus.

FILONÚS: Reconoces, pues, que no puedes concebir cómo una cosa corpórea sensible existe en otra forma que no sea en la mente.

HILAS: Así es.

FILONÚS: Y, sin embargo, has defendido seriamente la verdad de algo que no puedes ni siquiera concebir.

HILAS: Confieso que no sé qué pensar; no obstante, aún me quedan algunos escrúpulos.¹⁸

Es probable que Hilas tenga razón al conservar algunos escrúpulos, ya que el argumento de Filonús se ha ganado un buen número de críticas, e incluso cierto desprecio. Podemos emplear contra él parte de la artillería que habíamos desplegado contra el argumento ontológico, y preguntarnos si Filonús no estará tergiversando subrepticamente el sentido de expresiones como «en la mente». También podemos cuestionar la solidez de la conclusión de Filonús: por muy convencido que esté de que Hilas no puede concebir una casa o un árbol como no sea «en» una mente, parece que el argumento podría llevarnos del mismo modo a la aplastante conclusión de que Hilas no puede concebir una casa o un árbol como no sea en la mente del propio *Hilas*. Y esto es demasiado radical incluso para Berkeley.

A pesar de todo, cabe sospechar que hay un rastro de verdad en la postura de Filonús, como es habitual en el caso de los grandes filósofos. Tal vez nos sentiremos más próximos a ella si pensamos lo siguiente: supongamos que la intención de Hilas fuera defender que la idea realista de un objeto «independiente» de su forma de comprender las cosas tiene sentido para él. Se propone «abstraer» las contingencias relativas a su propia experiencia perceptiva y las contingencias relativas a su propia forma de pensar, es decir, sus propias preferencias entre conceptos. En tal caso Berkeley, a través del personaje de Filonús, le estaría recordando que tal abstracción es una hazaña imposible. Sea lo que sea aquello que imagine o conciba, vendrá marcado inevitablemente por su propia perspectiva.

Por ejemplo, es posible que Hilas imagine su árbol con un tronco marrón y unas hojas verdes. En tal caso, Berkeley tendría todo el derecho a alegar que esto no es cumplir la hazaña de imaginar un objeto más allá de la perspectiva humana, ya que los colores de las cosas son el resultado de tal perspectiva. La idea queda más clara en el caso de las cualidades secundarias, pero en este punto del primer *Diálogo* Berkeley ha preferido dejar que sea el lector quien la aplique

de forma más general. Un buen experimento mental para ilustrar esta postura sería el siguiente: supongamos que nos piden que imaginemos una habitación con un espejo en una pared y una mesa con un jarrón lleno de flores, situada en algún lugar frente al espejo. Nos avisan de que no debemos imaginarnos a nosotros mismos en la habitación. Estamos convencidos de poder hacerlo. Luego nos preguntan si el jarrón se refleja en el espejo. Si decimos «sí», nos hemos situado subrepticamente en una perspectiva determinada, y si decimos «no» nos hemos situado en otra, ya que el jarrón aparecerá en el espejo mirado desde ciertos ángulos y desde otros no. Difícilmente podremos responder «ni sí ni no», y tampoco podemos escaparnos del problema diciendo que a veces sí y a veces no, porque eso significaría mover nuestro punto de vista a lo largo de la habitación. Parece que estamos atrapados: el punto de vista aparece de forma imprevista cada vez que usamos la imaginación.

Berkeley nos está recordando la inevitable influencia que ejerce nuestra propia perspectiva sobre aquello que imaginamos o concebimos. Podemos darnos cuenta de la fuerza y la importancia que tiene esta idea si pensamos por un momento en un filósofo que la ignoró, a saber, G. E. Moore. Éste se propuso refutar la idea de que la belleza está en el ojo del espectador, o en otras palabras, se propuso defender una postura realista acerca de la belleza. Su argumento consistía en el experimento mental del «aislamiento».¹⁹ Moore nos pide que imaginemos dos mundos. Uno está lleno de nubes de algodón, árboles verdes, ríos saltarines y otras delicias pastorales. El otro es un montón de ceniza y basura. Luego especifica que en estos mundos no hay nadie. No son observados. Pero ¿acaso no es más bello el primero que el segundo? ¿Y acaso no demuestra eso que la belleza es independiente del ojo del espectador?

Filonús nos pone en guardia contra este insidioso argumento de Moore. Somos nosotros quienes aceptamos la invitación de pensar acerca de aquellos mundos. Y también somos nosotros quienes

aplicamos sobre ellos nuestras preferencias estéticas, que sin duda incluyen el gusto por el campo y el desagrado por la ceniza y la basura. Pero no nos hemos puesto por encima de estas preferencias ni las hemos dejado en suspenso al juzgar aquellos mundos imaginarios. Al contrario, son precisamente estas preferencias las que resuenan en nuestro veredicto. Todo lo que consigue demostrar Moore es que podemos considerar que algo es bello con independencia de si pensamos que alguien lo está viendo, y esto no refuta la concepción idealista o subjetivista de que la belleza se encuentra a pesar de todo en el ojo del espectador.

Me atrevería a decir que la mayoría de la gente encuentra sospechoso el argumento de Moore, y en esta medida simpatiza con Filonús. Pero la mayoría de la gente también encuentra sospechoso lo que dice Filonús. Si queremos refutar a Moore y a Filonús al mismo tiempo, tendremos que esforzarnos por encontrar un fundamento sólido. Tal vez Kant nos esté señalando el camino. Démosle la razón a Filonús cuando se trate de cosas como la belleza o las cualidades secundarias, pero tengamos claro que el «ojo del espectador» no tiene tanto peso cuando se trata de categorías más importantes del pensamiento, como por ejemplo las nociones de espacio, tiempo, los objetos físicos, el yo o la causalidad. Pero en la próxima sección daremos un rápido vistazo a algunos casos en los que incluso esta promesa de síntesis camina sobre un terreno inseguro.

Reglas y universales

La tradición idealista pone el acento en el papel crucial e ineludible que juega la configuración de nuestra propia mente *en* la «construcción» de nuestra concepción del mundo. Cuando se trata de señalar cuál es el factor determinante en la configuración de la mente, podemos escoger entre varias opciones. Berkeley y la

tradición empirista parten de la naturaleza subjetiva de la experiencia sensible, en especial la experiencia de las cualidades secundarias. Sin embargo, el hecho de que dependan del «ojo del espectador» las convierte en incómodas.

En la actualidad daríamos mayor importancia a factores culturales y sobre todo lingüísticos. No estamos tan preocupados por la subjetividad de la experiencia como por las variaciones culturales. En consecuencia, muchos pensadores contemporáneos se adhieren a una línea de pensamiento que tiene su punto de partida en Wittgenstein: la idea de las llamadas «reglas de uso».²⁰ Wittgenstein comienza por examinar el momento de la comprensión: cuando se nos explica cierto concepto y nosotros decimos: «Ahora ya sé seguir» o «Ahora ya sé lo que significa». Parece como si hubiéramos captado cierta regla o principio que distingue la aplicación correcta de un término de las aplicaciones incorrectas. Se trata de una auténtica hazaña. Algunas personas, y los animales en general, son demasiado estúpidos o distintos de nosotros como para comprender (hemos visto ya el ejemplo de los perros, que son incapaces de captar el significado del gesto de señalar algo y se dedican, en cambio, a mirar nuestro dedo). El hecho de que seamos capaces de realizar esta hazaña, y de que seamos capaces de repetirla en suficientes ocasiones, viene dado simplemente por la naturaleza. Es lo que hace posible la comunicación y los significados compartidos. Pero no podemos dar por supuesto que todos lo hacemos de la misma forma, o de una determinada forma en particular. Para ello sería necesario que nuestras mentes estuvieran configuradas del mismo modo. Pero ¿de qué depende la configuración de nuestra mente?

En realidad, se trata del antiguo problema de los universales, al que tantos esfuerzos dedicaron los griegos. Comprender las cosas y describirlas requiere el uso de conceptos que se ajustan a las reglas mínimas que acabamos de ver. Pero ¿cuál es el grado de «realidad» de estas reglas? En este punto se distinguen tradicionalmente tres

posibles posturas:

1.*Realismo* (a veces llamado *platonismo*). Estas reglas tienen una existencia real y objetiva. Determinan la correcta aplicación de nuestros conceptos sobre el pasado, el presente, el futuro y las instancias hipotéticas. Nuestra comprensión de estos conceptos depende de un misterioso acto de aprehensión que no resulta fácil de explicar en términos naturales.

2.*Conceptualismo*. Estas reglas son fruto de nuestra mente. Son conjeturas basadas en las respuestas que compartimos ante ciertos estímulos, debido a nuestra común naturaleza humana, o quizás a la cultura o a la educación. En este sentido todos nuestros conceptos serían dependientes de nuestras respuestas: instrumentos derivados de nuestra forma habitual de responder ante las cosas.

3.*Nominalismo*. Tales reglas no existen en realidad. No hay más que seres humanos con ciertas disposiciones a usar o no determinadas palabras. En realidad, ninguno de estos usos es «correcto» o «incorrecto», aunque a menudo aquellos usos que se apartan de los habituales son tachados de incorrectos.

Tal vez nos ayude pensar en un ejemplo que muestre las ventajas de cada postura. Pensemos en un concepto de dudosa validez, como el de «histeria» o «neurastenia». Ante un concepto así, un realista supondrá que existe un fenómeno real, cuyos límites están bien definidos (algunas de las personas que van por el mundo son histéricas y otras no). Al usar este término «cortamos la naturaleza siguiendo sus propias articulaciones», para emplear una metáfora desagradable. En cambio, un conceptualista rechazaría esta metáfora. No obstante, es posible que aceptara el concepto mismo de histeria, suponiendo que constituyera un principio o categoría útil para establecer el perfil de un determinado estado médico o psicológico. Tal concepto no haría más que agrupar una serie de

casos que nos parecen similares, que es todo lo que a fin de cuentas puede hacer un concepto. Finalmente, un nominalista diría que esta palabra es tan buena como cualquier otra. Si la gente está dispuesta a usarla, pues perfecto, ya que eso es todo cuanto se puede decir acerca del uso de cualquier término.

Naturalmente, estas posturas se presentan en diversas versiones, y cada una de ellas tendrá en principio sus ventajas y sus defectos. Para el realista o platonista, las otras opciones no hacen más que acercarnos al abismo idealista: la mente que construye o inventa su propia realidad (si la historia no es un fenómeno unitario en la realidad, no tiene sentido describir el mundo en función de ella. No nos permite alcanzar ninguna verdad objetiva). Tanto para el platonista como para el conceptualista, el nominalismo es completamente insostenible: sería tanto como negar la realidad del significado y de los conceptos y, por lo tanto, del pensamiento en general. Vendría a ser un «eliminacionismo», es decir, la negación del acto mismo de pensar. No hay ninguna mente que se limite a responder a los estímulos con meros exabruptos verbales, sin que nada gobierne la corrección o veracidad de las respuestas. En cambio, para el nominalista el platonismo resulta imposible de creer, y el conceptualismo no hace más que entregarse a la retórica de las reglas, sin entrar nunca en el fondo de la cuestión. Pues ¿qué diferencia hay entre unas reglas «constituidas» por nuestras disposiciones o respuestas y unas reglas que simplemente nos inventamos sobre la marcha? ¿Y qué diferencia hay entre esto último y no tener reglas en absoluto? Desde su punto de vista, un conceptualista no es más que un nominalista demasiado cobarde para reconocer que lo es.

Se dice que en la Edad Media los estudiantes llegaron a pegarse por las calles de París por culpa del problema de los universales. La cuestión es de vital importancia, ya que está en juego toda nuestra concepción de nuestra capacidad de describir el mundo de forma

verdadera o falsa, y la objetividad de cualesquiera opiniones que podamos tener. Incluso se podría decir que es el problema más profundo y fundamental al que se enfrenta la filosofía. Es el mismo problema que articuló el debate entre Platón (realista) y los sofistas (nominalistas). En realidad, lo que a menudo se llama «posmodernismo» no es más que nominalismo vistosamente presentado como la doctrina de que no hay nada más que textos. La misma clase de nominalismo paraliza cualquier apelación a la razón o a la verdad en muchos estudios de humanidades. La filosofía «analítica» hace el papel de Platón frente a sus sofismas, y trata de ahogar su atractivo canto de sirena.

En los últimos años, algunos filósofos han considerado razonable una versión «naturalizada» del realismo que permite evitar los misterios del platonismo. De acuerdo con ella, las cosas poseerían una serie de propiedades reales, con independencia de si las contemplamos desde una u otra perspectiva. Nuestras mentes estarían diseñadas precisamente en función de estas propiedades. El éxito en la evolución nos ha modelado para que seamos sensibles a estas clases reales que corresponden a la naturaleza de las cosas. Por mucha razón que tengan los conceptualistas al hacer hincapié en el carácter contingente de la configuración de nuestras mentes, cometen un error al olvidar que nuestras mentes no existen en el vacío. Han sido configuradas de forma natural por las estructuras causales del mundo en el que vivimos. Si las circunstancias son favorables, todos «vamos en la misma dirección» porque en el contexto del mundo tal es la dirección correcta. Un naturalismo de estas características, por ejemplo, conectaría con el resumen de la descripción científica de los colores que vimos en el capítulo 2. Su intención sería demostrar que incluso nuestra forma de clasificar las cualidades secundarias está lejos de ser arbitraria. Y si éstas ganan algo de estatus «realista», las otras cualidades habrían de seguir el mismo camino.

Se trata de una forma bastante confortable de ver las cosas, y se ajusta perfectamente al «fundacionalismo natural» que vimos al final del capítulo 1: la defensa de la armonía entre la mente y el mundo inspirada en el evolucionismo. Cabe tener la esperanza de que sobreviva a la batalla intelectual que he intentado hacer que resonara en este capítulo. Pero antes de considerar que se han terminado todos nuestros problemas, es necesario que estemos convencidos de que la imagen científica o absoluta del mundo está en perfecto acuerdo con su imagen manifiesta. Tendríamos que poder pensar que Kant o algún sucesor suyo ha conseguido navegar con éxito entre Filonús y Moore, o bien que ha conseguido resolver el problema de Hume con la «filosofía moderna» sin realizar demasiadas concesiones al idealismo. Y no todo el mundo está convencido de eso.

CAPÍTULO

8

¿Qué hacer?

Hasta ahora nos hemos interesado por nuestra comprensión del mundo. Nos hemos ocupado de la naturaleza de las cosas, de nuestro conocimiento de ellas y de las diversas formas de razonar acerca de ellas. Pero buena parte de nuestros razonamientos no son teóricos, no se interesan por aquello que el mundo pueda ser, sino que son prácticos, es decir, se interesan por la forma en que debemos actuar en él. Pensamos acerca de lo que debemos hacer y repasamos los argumentos y consideraciones que tiene a su favor cada posible curso de acción. ¿Qué debemos pensar acerca de todo esto? Hay tratados enteros y enciclopedias consagrados a este tema, cuya esencia serían la ética y la filosofía moral, aunque no lo agotan, ya que los razonamientos prácticos no tienen siempre, ni mucho menos, una naturaleza exclusivamente moral. Hay también cuestiones técnicas y estéticas que tratar junto a los problemas morales. En este último capítulo no pretendo cubrir todas las materias que aparecen en los tratados. Resulta imposible hacerlo en tan poco espacio. Pero creo que se pueden encontrar algunas ideas sólidas en este campo, y mi intención es sugerir cuáles son.

Inquietudes reales

Buena parte del pensamiento práctico tiene una naturaleza técnica. Tenemos un objetivo, y nuestro problema consiste en encontrar el modo de alcanzarlo. Nos esforzamos por ajustar los medios a los fines, partiendo siempre de los fines. El fin es algo dado: queremos arreglar la nevera o plantar flores o construir un puente. Evidentemente, podemos estar más o menos capacitados para hacer estas cosas. No hay una única «manera de pensar» que nos permita alcanzar todos nuestros objetivos, por la misma razón que la persona que sabe arreglar la nevera no tiene por qué saber plantar las flores o construir el puente. Adquirir las habilidades necesarias requiere comprender los elementos que intervienen, y saber qué cambios hay que introducir y cómo hay que hacerlos para obtener el resultado deseado.

Se dice habitualmente que nuestros objetivos vienen marcados por nuestros deseos, de modo que el pensamiento técnico tendría por objeto satisfacer de forma eficiente nuestros deseos. Y muchas veces es cierto, al menos de forma aproximada. Pero pensar de este modo podría llevarnos a confusiones. Si concebimos el deseo como un estado de entusiasmo por un determinado fin —cuando algo nos hace brillar los ojos—, hay que decir que muchas veces actuamos movidos por ciertas inquietudes que no quedan bien expresadas con la palabra deseo. Aquí estoy cortando el césped, cuando me gustaría estar navegando. ¿Por qué? No porque desee cortar el césped. Tal vez deteste hacerlo. Pero había llegado el momento, es decir, tenía que hacerse. Tengo interés en que el césped esté cortado, y adopto las medidas necesarias para conseguir este objetivo. Tener una cierta inquietud significa en este caso que el pensamiento de que el césped necesita ser cortado me incita a actuar. Puede ser que piense que me corresponde a mí cortar el césped. O bien puede ser que piense simplemente: «Ya es hora de hacerlo», sin pensar conscientemente

en el papel que me corresponde en el cuidado de la casa o lo que sea. Sin embargo, es frecuente que piense que el césped de otra persona necesita ser cortado, sin sentir ningún impulso de cortarlo yo mismo. De modo que es el interés que tengo en el cuidado de mi casa el que me ha hecho sensible al pensamiento de que mi césped necesita ser cortado, incluso sin reflexionar abiertamente sobre este papel que me atribuyo.

Hay una diferencia importante entre actuar movidos por alguna inquietud y actuar simplemente porque queremos. A menudo, esta diferencia se ignora deliberadamente en las discusiones. Imaginemos una relación que pasa por ciertas dificultades. Annie se siente obligada a abandonar a Bertie por cierto motivo: tal vez algún compromiso con otras personas o un proyecto de vida que requiere un cambio de residencia. Bertie puede aumentar la temperatura emocional del debate a base de insistir en que Annie no se iría si no quisiera hacerlo. «Lo que pasa es quieres irte, de otro modo no lo harías.» Tales palabras resultan hirientes, ya que vienen a decir que la posibilidad de abandonar a Bertie hace brillar los ojos de Annie, o que ella valora este hecho como un aspecto positivo de su forma de actuar. Y esto podría ser completamente injusto. Annie podría estar completamente deshecha ante la idea de abandonar a Bertie y, sin embargo, verse obligada a hacerlo, igual que sucedía con el césped.

Alguien podría sugerir que cuando actuamos movidos por alguna inquietud, lo que nos empuja en realidad es nuestro deseo de algo relacionado con el resultado. Cuando digo que tengo interés en cortar el césped, aunque no me guste hacerlo, hay que pensar que mi acción responde a un deseo de alguna otra cosa: quizá sólo una cierta tranquilidad de espíritu, por ejemplo. Nos encontramos aquí con otro error peligroso, que consiste en pensar que siempre que una persona se compromete con algo, lo que «realmente» desea es alcanzar cierto estado de ánimo, como por ejemplo su propia tranquilidad de espíritu. A menudo los psicólogos han tenido tendencia a presentar el deseo

como una especie de aumento de la tensión que el agente se vería obligado a liberar. En tal caso, resulta fácil llegar a la conclusión de que el verdadero objeto del deseo ha sido desde el principio liberar la tensión. Esto es algo que también puede provocar comentarios hirientes: «No es que estés realmente comprometido con el problema de los niños hambrientos, lo único que querías era sentirte bien». Todo comportamiento se considera esencialmente egoísta, como si lo único que nos preocupara fuera nuestra propia situación, y los demás objetivos e intenciones no fueran más que una máscara.

Esta manera de pensar (a veces llamada «egoísmo psicológico») está completamente equivocada.¹ Supongamos que me encuentro con una persona que quiere comer. Siguiendo la línea de razonamiento del último párrafo, interpreto que quiere liberar la tensión provocada por su deseo de comer. De modo que le doy un puñetazo en el estómago, para que el dolor le haga olvidar la comida. ¿Acaso le he dado lo que quería? En absoluto (incluso si olvidamos el hecho de que el puñetazo debe de haber sido doloroso). Él no quería que le liberara de ninguna tensión. Lo que quería era comida. Asimismo, una persona normal que sienta un deseo sexual no busca liberar su tensión. El bromuro podría tener ese efecto, pero no es eso lo que quiere. Lo que quiere es sexo.

Pasemos ahora a considerar otro tipo de inquietudes. Supongamos que soy un padrino de la mafia y que estoy convencido de que Luigi me ha insultado. Ordeno a uno de mis secuaces que se deshaga de Luigi. Mi hombre se va, algo intimidado por la peligrosa tarea que le he encomendado. Sin embargo, razona él, lo que yo quiero realmente es liberar la tensión que me provoca la existencia de Luigi. Y él puede liberarme de ella de otro modo: organizar un engaño perfecto que me haga pensar que ha matado a Luigi. Así que eso es lo que hace, mediante un montaje convincente. ¿Acaso ha hecho lo que yo quería? Resulta claro que no. Yo no quería vivir en un engaño que me hiciera pensar falsamente que Luigi estaba muerto (mejor no pensar en lo

que sucedería si descubriera la situación en la que me había puesto mi propio matón). Lo que yo quería era que Luigi estuviera muerto.

La conclusión podría ser: una de nuestras inquietudes consiste precisamente en evitar que nos engañen acerca de ellas.

De nuevo descubrimos un importante motivo de controversias y malentendidos. A menudo, la comunicación depende de comprender las inquietudes de los demás, y ello no es posible si contemplamos estas inquietudes como un problema o un obstáculo en sí mismo: algo que debe ser manejado o curado. Supongamos que Annie estuviera preocupada por su carrera y su desarrollo como persona, y Bertie reaccionara buscando formas de disipar esta preocupación, en lugar de apoyarla en su carrera y en su desarrollo como persona. «Cariño, no te pongas de mal humor. Ya verás como todo esto dejará de preocuparte si salimos a cenar/nos damos la mano/tenemos un hijo...» La respuesta es inadecuada en el mismo sentido en que lo era el puñetazo en el estómago para aliviar el hambre. Pero es probable que este hecho no resulte tan evidente, por lo menos para Bertie, y es probable que siguiera sin comprenderlo por mucho que Annie rompiera con él. Para emplear un término que ya introduje en el capítulo 3, podemos decir que Bertie ha «objetivado» la preocupación de Annie, como si este mismo hecho fuera el problema, en lugar de fijarse en aquello que tenía a Annie preocupada. Pero desde el punto de vista de Annie, el problema es su carrera, no su preocupación por su carrera. Mientras Bertie siga sin compartir esta forma de ver las cosas, no se pondrán de acuerdo.

Esta cuestión tiene importantes repercusiones para la industria y la cultura de la «terapia». Volveré a este tema en cuanto haya añadido un par de puntos más a mi esquema.

Tal como dije antes, una de nuestras inquietudes consiste precisamente en evitar que nos engañen acerca de ellas. Una idea parecida es que a menudo, aunque no siempre, una de nuestras inquietudes es precisamente no abandonar nuestras inquietudes.

Supongamos que le dicen al padrino que quería ver muerto a Luigi que si se espera diez años, se le va a pasar el deseo («De modo que al final todo estará solucionado», podría decir alguien). Esto sería lo mismo que decirle a la pareja preocupada por su carrera que si espera a tener un niño, su preocupación va a disminuir. Uno no quiere abandonar sus inquietudes. Lo mismo se puede expresar diciendo que el padrino se identifica con su deseo de venganza y que la mujer se identifica con su preocupación por su carrera.

Ahora bien, hay casos en los que no nos identificamos con nuestros deseos e inquietudes. Nos gustaría perderlos de vista como fuera. Cuando alguien siente el deseo irreprimible de fumar un cigarrillo, es posible que no sólo desee el cigarrillo, sino también deshacerse como sea de este deseo. Una buena solución sería la terapia o algo así como una extracción quirúrgica del estado mental en cuestión. Si estamos «obsesionados» por alguien o por algo, es posible que pensemos que sería mejor deshacerse de la obsesión, e incluso es posible que tomemos medidas para lograrlo. Reconocer un deseo con el que nos hemos identificado como una obsesión o una necesidad incontrolada es ya una forma de distanciarnos, y el principio del proceso de objetivación que habrá de permitirnos encontrar alguna estrategia para escapar de él. La esposa preocupada por su carrera que hemos visto en el ejemplo anterior podría acabar por compartir la idea de su marido de que el problema es precisamente su ambición, y buscar con mayor entusiasmo otras distracciones que le ayudaran a olvidarla. Pero también podría ser que no lo hiciera, y tal vez habría sido un error olvidar la preocupación por su carrera, ya que podría ser más importante para su identidad de lo que había pensado.

Todo esto demuestra que no siempre resulta evidente la diferencia entre aquellas inquietudes con las que nos identificamos y aquellas inquietudes que podemos objetivar. Tal vez tengamos que ponerlas a prueba para saber si es posible (o adecuado) abandonar ciertas

inquietudes, o si lo único que podemos hacer, o quizás lo más adecuado, es seguir adelante con ellas y tratar de alcanzar nuestros objetivos.

¿Qué significa tener una inquietud? Antes dije que es actuar movido por una idea. Hay un aspecto de la realidad que nos empuja a actuar y que se convierte en un elemento que tiene cierto peso para nosotros, o que nos importa (es interesante señalar que las metáforas que surgen espontáneamente se refieren al peso o a la presión). Cada uno de los aspectos de la realidad tiene su peso en nuestras decisiones, evidentemente. Una de las formas que tienen de pesar sobre nosotros es provocar ciertas actitudes, como la admiración o el desprecio, o ciertas emociones, como el miedo o la esperanza. Cuando leo una obra de ficción, por ejemplo, es posible que me disguste un determinado personaje, lo que significa que los términos que se emplean para describirlo tienen su peso o su influencia sobre mí. Me siento empujado a admirar las virtudes del héroe o a detestar los vicios del villano.

Cuando nos mueve una determinada inquietud, se puede decir que aquellos aspectos de la realidad a los que somos sensibles son las razones por las cuales elegimos una cosa en lugar de otra, adoptamos una determinada actitud o sentimos una determinada emoción. La razón de que corte el césped es que le hacía falta. La razón de que Annie abandone a Bertie es que su carrera exige que cambie de residencia. En este sentido, nuestras razones son aquellos aspectos de una situación que influyen en nuestras deliberaciones acerca de lo que debemos hacer o acerca de cómo debemos sentirnos ante cierto hecho. En un sentido un poco más amplio, nuestras razones pueden ir más allá de lo que nos planteamos de forma explícita en nuestras deliberaciones. Pueden incluir aspectos de la situación que tienen una influencia sobre nosotros, aunque no nos demos cuenta de ella, o sólo a medias. En este sentido más amplio, la razón de que Annie abandone a Bertie podría ser que está aburrida de él, aunque no lo

admita abiertamente ante sí misma.

Hay que señalar una distinción importante en nuestra forma de hablar acerca de las razones que mueven a las demás personas. Podemos hablar de forma descriptiva o normativa, es decir, podemos describir aquel aspecto de la situación que les empuja a actuar, o bien podemos dar nuestra opinión acerca de si eso que les preocupa es o no es realmente una razón, con lo que expresamos nuestro acuerdo o nuestro rechazo por sus razones. Es importante tener presente esta distinción. Si decimos que Annie no tenía ninguna razón para abandonar a Bertie, es posible que pretendamos hacer una observación (probablemente falsa) acerca de la psicología de Annie: que actuó simplemente por impulso, de forma irreflexiva y sin obedecer a ningún deseo o interés en particular. O bien, de forma más razonable, es posible que pretendamos *rechazar las* razones que movieron a Annie a actuar: se fue porque estaba interesada en continuar su carrera como bailarina de ballet, pero en aquellas circunstancias eso no era más que una necia ambición, o algo que no deber/a haber pesado en su decisión. Cuando hablamos de forma normativa deberíamos indicarlo mediante el uso de palabras como «debería» y «bueno». Pero en lugar de decir: «No tenía ninguna buena razón para hacerlo», a veces decimos cosas como: «No tenía ninguna razón para hacerlo», lo que puede llevar a confusiones.

Según parece, podemos tener las inquietudes más dispares. La muerte de un enemigo, el progreso en nuestra carrera, el estado del césped, el bienestar de nuestra familia y amigos, son inquietudes o intereses muy comunes, al igual que tantos otros: el hecho de haber realizado una promesa, el hecho de que alguien nos hubiera hecho un favor, el hecho de que seamos la esposa de alguien, o un médico o un abogado. Hay tantas inquietudes distintas como personas y tipos de personas. Y de momento ya hemos rechazado un intento de reducir esta diversidad a cierta unidad: el intento de reducir nuestras motivaciones a la búsqueda de determinados estados mentales

(liberar las tensiones provocadas por nuestros deseos y preocupaciones). Pero se trataba de un error, ya que pasaba por encima de la distinción entre inquietudes con las que nos identificamos e inquietudes de las que podemos distanciarnos e intentar perder de vista.

La voz interior

Muchas inquietudes son privadas y opcionales. Supongamos que me interesan los trenes de vapor. En tal caso, el hecho de que en cierto lugar pasaran los trenes tendría su peso para mí. Desde mi punto de vista, sería una buena razón para ir a ese lugar. En cambio, no tiene por qué tener la menor importancia para ti. Y tampoco tiene por qué importarme a mí si a ti te interesan o no los trenes. Incluso podría alegrarme de ello, ya que la vista es mucho mejor si hay poca gente.

Sin embargo, hay otras inquietudes que esperamos encontrar en las demás personas. Es decir, que una de nuestras inquietudes es que las demás personas se preocupen por ciertas cosas. Hay cosas que esperamos que influyan en sus decisiones y actitudes: el hecho de que ciertas acciones puedan defraudar a otras personas, o que signifiquen la ruptura de una promesa, o bien que constituyan un comportamiento deshonesto o manipulador, entre otras muchas. Del mismo modo, esperamos que ciertas acciones causen un cierto pesar en las demás personas. Nos sorprendería e incluso nos dejaría perplejos que no fuera así. Esto nos lleva al terreno tradicional de la ética. ¿Cuáles son las inquietudes que podemos esperar de los demás?

Podemos distinguir dos formas posibles de abordar esta cuestión. La primera consiste en preguntar cuáles son las inquietudes que definen una vida ideal. ¿Cómo debemos vivir? Las diferentes tradiciones éticas han respondido a esta pregunta de maneras

distintas. La vida ideal de un héroe homérico está llena de inquietudes en torno a su honor, su condición y su éxito en la batalla. La vida ideal para un santo cristiano está repleta de inquietudes que incluyen el amor a Dios, la supresión del orgullo y un conjunto de ideales sobre el amor fraternal. Según el confucianismo, la vida ideal contiene una importante dosis de respeto por los códigos tradicionales. Todos estos ideales pueden materializarse de una u otra forma o pueden ser pintados con unos colores más o menos atractivos. Y, sin embargo, hay algo en ellos que nos sigue resultando incómodo, aunque sólo sea porque tenemos escasos motivos para pensar que existe algo parecido a la vida ideal. Desde el momento en que personas distintas tienen gustos e intereses distintos, y que las diferentes culturas alientan inquietudes distintas, parece razonable pensar que toda «vida ideal» estará fuertemente contextualizada: ideal para determinada persona en determinadas circunstancias, pero no mucho más. Ni siquiera los ingredientes de una buena vida, por no hablar de una vida «ideal», resultan tan evidentes. Algunos componentes esenciales suscitan poca controversia. La mayoría de la gente incluiría la salud (y los medios necesarios para asegurarla), la felicidad (en su sentido auténtico: no la que resulta de vivir en un paraíso de bobos), los logros (de nuevo en el buen sentido: no la realización de vanas o estúpidas ambiciones), la dignidad, los amigos, el amor, la familia. Más allá de este tipo de cosas, asuntos como la riqueza o el tiempo libre podrían ser controvertidos, e incluso algunas variedades de los componentes esenciales podrían ser vistas antes como un castigo que como una bendición. Ciertas personas podrían haber disfrutado de una vida mejor si, pongamos por caso, no hubieran sido bendecidas con una salud de hierro que les hiciera incapaces de sentir compasión ante las debilidades de los demás.

Sin embargo, si interpretamos la pregunta por las inquietudes que debemos esperar de los demás en otro sentido, parece que podemos llegar a conclusiones más estrictas. De acuerdo con esta

interpretación, se trata de preguntar por los límites justos de la conducta. Y es en este sentido en el que podemos decir que cuando no respondemos a las expectativas, hemos hecho algo mal. No hemos estado a la altura y, por lo tanto, nos hemos hecho merecedores de diversos tipos de reproches. La gente espera de las demás personas que sean honestas, cooperativas, sensibles a las necesidades del otro, justas, bienintencionadas y otras cosas por el estilo, y si fallamos en alguno de estos apartados se puede decir que no hemos estado a la altura y es posible que seamos censurados. Las otras personas tienen quejas acerca de nosotros; les preocupa que seamos de esta manera.

Alguien podría sentirse irritado por todo esto. Tal vez le quitar/a importancia a la opinión negativa de los demás. ¿Por qué le tendría que inquietar? ¿Por qué no ser un espíritu libre, alegremente despreocupado de lo que piense el mundo? En algunos casos hay algo admirable en esta actitud: el visionario y el santo pueden parecer despreocupados de la opinión del mundo mientras se esfuerzan por cambiarlo, quizás para bien. Pero la pregunta es: ¿por qué nos estamos ganando la mala opinión de los demás? Si nos la estamos ganando, por ejemplo, porque nos importa un pepino respetar las promesas o no tenemos reparo en meter mano al dinero ajeno, entonces será difícil librarnos de su censura. Actuar de este modo — ser capaces de mirarles a los ojos y decirles que no vemos de qué se quejan— requiere no sólo una falta total de respeto por las promesas o por la honestidad, sino también una falta total de reconocimiento hacia las inquietudes de los demás en lo que respecta a estas cosas. Entre la gente corriente, raras veces se encuentra un grado tal de insensibilidad. Una cosa es el típico rufián que espeta: «No me importa si te he hecho daño al romper mi palabra o al robar tus bienes». Pero otra muy distinta es alcanzar la cota extraordinaria de villanía que supone decir: «Ni siquiera reconozco que te puedas quejar». Suele ser más fácil decir esto a modo de desafío arrogante

que sentirse cómodo en esta actitud, aunque también es cierto que la moralidad sexual promueve en ciertas ocasiones que la persona culpable de la mala acción sea incapaz de ver el motivo —lo cual no hace más que empeorar las cosas—. Una sociedad donde la gente fuera incapaz de reconocer en ningún caso las quejas de los demás sería una sociedad sin ética, aunque por la misma razón sería difícil reconocerla como sociedad.

Los pensadores han intentado articular estas ideas de varias maneras distintas. «Interiorizar» un conjunto de valores es algo muy parecido a interiorizar la mirada o la voz de los demás. Reconocer que tienen derecho a quejarse de nosotros equivale a ver que uno mismo ha fallado a ojos de los demás, e interiorizar su voz significa que uno mismo cargue con esta idea. El malestar se traduce en autorreproches o en sentimientos como la vergüenza y la culpabilidad. Muchos sistemas éticos se basan en alguna versión de la conocida Regla de Oro: «Trata a las demás personas tal como desearías que te trataran a ti». Algunos pensadores insisten en la emergencia de un «punto de vista común»; otros, en la simpatía o empatía que permite que la forma en que nos ven las demás personas resuene en nuestra forma de vernos a nosotros mismos. Hume ofrece un espléndido ejemplo para mostrar la facilidad y la naturalidad con la que incorporamos la visión de los demás a nuestras inquietudes: «Una persona puede sentirse ofendida si se le dice que le huele el aliento, a pesar de que ello no le cause la menor molestia física».² Nos vemos a nosotros mismos desde el punto de vista de los demás y el resultado puede provocarnos malestar o satisfacción.

Podemos describir este aspecto de nuestra psicología en términos de una adopción de las razones del otro. Si tienes un piano sobre el pie, una de tus inquietudes será cambiarlo de sitio lo antes posible. Si yo soy consciente de esta situación, lo natural será que comparta esa inquietud, y en caso contrario estaría faltando a las expectativas. No

ocupo el mismo lugar en la situación, ya que el piano no me hace daño a mí, sino a ti. Pero de mí se espera que simpatice contigo, que adopte tu inquietud, que ayude, y que trate tu problema como si también fuera el mío. Lo que para ti es una razón para actuar, para mí es una razón para ayudar. A algunos filósofos morales les gusta pensar que en estos casos existe algún tipo de imperativo de la razón. Piensan que habría algo defectuoso en mi racionalidad, o en mi forma de entender las cosas, si no adoptara tu inquietud y la convirtiera en mía. Yo no aconsejo esta manera de verlo. La persona que se muestra indiferente ante esta situación es mala, sin duda. Y puede haber cosas que no funcionen en su razonamiento, o en su manera de entender el mundo. Puede tratarse de un psicópata, incapaz de comprender la realidad de los demás. O tal vez realice un análisis deficiente de la situación y piense que lo que en realidad nos conviene es sufrir. Pero el caso más común es que el otro desvíe la mirada, o bien pase por otro lado, y no hay razón para pensar que haya ningún problema en su forma de entender el mundo o de razonar acerca de él. Es duro de corazón, no de mollera. Lo primero es tan malo como lo segundo, o incluso peor. Y cuando situamos el defecto en el lugar adecuado, vemos que no necesita tanto mejorar la estructura de sus razonamientos como la educación de sus sentimientos.

Verdad y bondad

Sin embargo, hay una cuestión que divide a los pensadores en dos bandos.

Consideremos la siguiente ecuación:

Una de las inquietudes de X consiste en perseguir/promocionar/apoyar $\phi = X$ piensa que ϕ es bueno/piensa que ϕ es una razón para actuar.

La división se produce entre aquellos pensadores que leen la ecuación «de izquierda a derecha», y aquellos que la leen de «derecha a izquierda». Es decir, algunos pensadores consideran que la explicación adecuada de nuestras motivaciones parte de las inquietudes, entendidas tal como hemos visto, para llegar finalmente a la «idea de que algo es una razón», la cual se justifica a partir de lo anterior. Por otro lado, están los que piensan que la explicación debe partir de la idea de que algo es una razón, tomada como una simple creencia acerca del caso en cuestión, para llegar hasta las inquietudes, las cuales se justifican a partir de lo anterior.

Estos dos planteamientos en el campo de la teoría ética reciben a veces el nombre de «no cognitivismo» y «cognitivismo». La idea es que si la ecuación se lee de izquierda a derecha, decir que algo es bueno, o que algo es una razón para actuar, constituye una reflexión a partir de ciertas actitudes mentales: el hecho de que algo influye sobre nosotros o que tiene cierto peso. Esta actitud mental no es una simple creencia. No es una representación de algún aspecto del mundo. Es una reacción ante cierta representación de los hechos. No se fija en ningún elemento en particular. En sentido estricto, por lo tanto, no es un estado mental susceptible de ser verdadero o falso, del mismo modo que tener ganas de tomar un café no es verdadero ni falso. La dirección no cognitivista queda bellamente expresada en palabras de san Agustín:

Existe la atracción de la voluntad y del amor, en la cual se manifiesta el valor de todo aquello que debe perseguirse, o evitarse, o tenerse en mayor o menor estima.³

Si la ecuación se lee en el otro sentido, de derecha a izquierda, vemos que el fundamento es una creencia: la creencia de que ϕ es una razón para la acción. Se trata de un tipo especial de creencia, ya que distingue o representa razones, pero es una creencia que genera inquietudes. Se suele decir que Aristóteles concebía de este modo las

motivaciones: su eslogan era que desear algo significa verlo como algo bueno. Es como si el deseo respondiera a una verdad percibida.

Esta cuestión es importante para muchos pensadores, en especial para los del bando cognitivista. Temen que si se pierde el punto de apoyo que aporta el cognitismo, el razonamiento práctico quedaría reducido a «meras» inquietudes, deseos y actitudes. En cambio, si podemos llevarlo de algún modo al terreno de la Verdad, tendremos algo en lo que fundar las pretensiones de la ética. Las inquietudes que corresponden directamente a estas verdades son las buenas; merecen tener autoridad sobre las demás.

Personalmente, creo que éste es uno de aquellos casos en los que toda la ventaja corresponde a un mismo bando: el no cognitivo.

La razón principal es que tiene que haber algo más que creencias o cogniciones, es decir, representaciones de aspectos de las cosas. También hay «la atracción de la voluntad y del amor». Cuando una persona tiene una inquietud, hay un aspecto de la situación que tiene cierto peso para ella. Este peso se traduce en una motivación más o menos fuerte: en un impulso para cambiar sus acciones y sus actitudes. ¿Acaso la «idea de que algo es una razón» puede llegar a tener este peso?

Hay varias sugerencias posibles acerca de cuál es el objeto de esta cognición. Una de ellas sería decir que se trata de un hecho puramente natural. Por ejemplo, la frase: «Ver que tienes el piano sobre el pie es una razón para sacártelo de encima» podría traducirse como: «Ver que tienes el piano sobre el pie y te está haciendo daño». Pero el problema es que es obviamente *contingente* si esto tiene algún peso para el agente. Si tiene el corazón de hielo o es un enemigo o tiene un peculiar sentido del humor, tal vez no le afecte en absoluto. De modo que tener una motivación o una inquietud no consiste en eso, ya que éstas influyen por definición sobre el agente. G. E. Moore lo resumió diciendo que sean cuales sean los rasgos naturales que reconocemos en las cosas, la forma en que se

relacionan con las razones de nuestras acciones sigue siendo una «cuestión abierta».⁴ Considerar que esta relación existe es dar un paso significativo: el paso que nos lleva al terreno de la práctica.

Otra posibilidad sería sugerir que el objeto de la cognición es un hecho peculiar, no natural, sino «normativo». Ésta era la perspectiva de Moore, y es posible que también fuera la de Platón.⁵ Es como si vislumbráramos algo distinto de los rasgos empíricos o científicos que conocemos en las cosas. Es como si vislumbráramos un orden normativo.

Todo esto suena muy misterioso. La ecuación leída de derecha a izquierda, si esto es lo que hay en el lado derecho, resulta bastante extraña. Supongamos que concebimos el mencionado orden normativo sobre el modelo de las leyes humanas. Así pues, es como si hubiéramos encontrado una ley que dijera que se debían sacar los pianos de encima de los pies. El problema es que nuestra actitud hacia una ley, como hacia cualquier otra cosa, depende siempre de nosotros. En principio, yo podría ignorarla perfectamente. Podría rechazarla de plano. No hay ninguna relación necesaria entre la existencia de una ley y el hecho de que tenga alguna influencia sobre mí. Así pues, no está claro que en esta dirección vayamos a encontrar algún tipo de explicación. Ocurre lo mismo, por cierto, cuando la ley viene marcada como «Ley de Dios». Esto es algo que podría no tener ninguna importancia para mí. Si así fuera, el arma tradicional que debería derrotarme sería el temor a la cólera de Dios. Pero una cognitivista no estaría dispuesto a apelar a un estado emocional contingente de este estilo, ya que eso sería sacar la cuestión del terreno de la razón. Lo que ella pretende es que el objeto de nuestra cognición nos motive necesariamente, que sea magnético.

Enfrentado a esto, un cognitivista sería presa del pánico. Respondería negando la ecuación con la que hemos empezado. Diría: «Bueno, acepto que existe un vacío entre percibir verdaderamente el orden normativo y estar motivado. Pero eso no es ningún problema:

cuando se tiene buena voluntad o buen corazón se hace lo que se ve que es razonable». Lo considero un acceso de pánico porque con esta respuesta el cognitivista puede proteger su querido compromiso con la idea de la Verdad, pero sólo al precio de ir a buscar la motivación fuera del dominio de la verdad. Desde esta perspectiva, el problema de las personas desprovistas de buena voluntad o buen corazón no sería que vieran verdades equivocadas. Sin embargo, todo el proyecto del cognitivismo consistía en llevar el razonamiento práctico al terreno de la verdad, lo cual nos permitiría decir que una persona con malas intenciones o con unas intenciones equivocadas está cometiendo un error desde el punto de vista de la razón, está interpretando mal el mundo. Si el cognitivista no puede seguir defendiendo esto al final de la discusión, no le ha servido de nada ganar alguna escaramuza particular a cambio de rendir su tesis.

En mi opinión, todos estos problemas desaparecen si leemos la ecuación en el otro sentido. Cuando la gente tiene inquietudes, se expresa hablando de razones, y considerando que las cosas que les importan son deseables o buenas. Lo hacen desde «la atracción de la voluntad y del amor». Creo que inventamos las proposiciones normativas («Esto es bueno», «Esto es una razón para la acción», «Deberías hacer esto») para poder reflexionar acerca de las inquietudes que nos exigimos a nosotros mismos y a los demás. Hablamos en estos términos para poner en claro nuestras motivaciones, para exponerlas a la admiración ajena, a la crítica y a la mejora. No estamos vinculados a ningún misterioso orden normativo.

Entonces, ¿no hay ningún conjunto de inquietudes que sea mejor que otro? Por supuesto que sí. Pero su superioridad reside en la forma de vivir que alientan. Un conjunto de inquietudes que nos conduzca a una vida leal, amistosa, agradecida, prudente, alegre y justa es claramente superior a uno que conduzca a vidas intrigantes, susceptibles, maliciosas, descuidadas, insensibles e injustas. Nuestras

vidas van mejor cuando merecemos la primera descripción, en lugar de la segunda. Y debería ser una de nuestras inquietudes que nuestras vidas fueran mejores.

Buenos malos sentimientos

Muchos escritos sobre ética plantean la cuestión de una forma distinta: introducen un dualismo. Por un lado tenemos los caóticos impulsos del deseo. Por el otro, por encima y a mucha distancia, tenemos los aristocráticos principios de la ética, que existen para controlar el deseo. Creo que con este cuadro sólo se consigue aumentar la confusión. Los aristocráticos principios de la ética resultan completamente misteriosos: parecen reclamar un origen divino o algún tipo de habilidad platónica para que se mantengan en armonía con la naturaleza de las cosas. Prefiero un modelo que no reconozca más que una pluralidad de inquietudes. No obstante, algunas de estas inquietudes tienen algo así como un estatus que nos lleva a hablar de virtud y vicio, deber y obligación. Son éstas las inquietudes que esperamos en los demás, de modo que si no las compartimos, o no tienen la suficiente importancia para nosotros, se considera que no estamos a la altura. Podríamos decir que es lo que las personas se deben unas a otras. Si alguien me hace un favor, le debo un sentimiento de gratitud: es su derecho y es mi deber sentirlo o expresarlo así. Si soy insensible o poco atento, no he estado a la altura. Perderé la admiración de los demás, y en la medida en que tengo una voz interior que se hace eco de la voz de los demás, me sentiré mal conmigo mismo. Si no me siento mal, esto mismo podría ser motivo de censura, a veces incluso más dura que la merecida por el fallo original. Si alguien no cumple con un deber de gratitud, podemos decir que se ha comportado mal. Pero si cuando le señalan su mal comportamiento, ella le quita importancia o no comprende el motivo de tanta queja, esta actitud puede resultar aún más

incomprensible que la falta original. Es por ello que concedemos tanta importancia a la constricción y, en casos graves, al arrepentimiento. Estos malos sentimientos son buenos.

En este punto podríamos retomar la crítica que habíamos expresado antes hacia la obsesión contemporánea por la «terapia». En nuestro ejemplo, la inquietud de Annie era su carrera, y esta inquietud no era apoyada o compartida por Bertie, que la veía como un problema en sí misma. Los casos morales son similares. Sentirnos mal con nosotros mismos o con nuestra conducta es ciertamente desagradable. Quisiéramos alejar estos malos sentimientos. Pero cuando están justificados, desear olvidarlos implica una cierta autoalienación, y no es la reacción adecuada. Supongamos que Annie sabe que ha hecho daño o que ha insultado a Bertie. Podría agradecer la ayuda de un terapeuta que le ofreciera un sencillo tratamiento para eliminar todo su sentimiento de culpa. Pero no está claro si debería agradecersele. Ante todo, su preocupación principal es arreglar las cosas con Bertie; pedirle disculpas o compensarle de algún modo, o asegurarle que sigue siendo alguien importante para ella, y ese tipo de cosas. O tal vez su inquietud podría dirigirse contra su propio carácter o su conducta, que desearía mejorar. Pero a Annie no le preocupan sus inquietudes en sí mismas. Y si un terapeuta pudiera darle una pastilla que se las quitara de encima, tal vez no la estaría ayudando. Ya no arreglaría las cosas con Bertie, ni se esforzaría por mejorar su propio carácter. En realidad, estaría convirtiendo a Annie en el tipo de persona que atrae grandes dosis de censura, no sólo por actuar mal, sino también por ser incapaz de reconocerlo tal como estaba haciendo ella. Está alienando a Annie de su conciencia de lo que ha hecho, y sería mejor que no lo hiciera.

Por supuesto, el sentimiento de culpa puede llegar a convertirse en un problema si dura demasiado tiempo o si se tiene mala suerte. Si éste fuera el caso, Annie haría bien en intentar librarse del autorreproche, ya que se habría convertido en una obsesión que no

tendría ningún efecto positivo. Pero lo cierto es que éste no es el caso más típico o fácil de resolver. Es un caso en el que las cosas se nos han escapado de las manos. Cuando las cosas están bajo control, el problema no es la culpabilidad o la vergüenza, sino las acciones que las han propiciado.

Nuestras inquietudes influyen sobre nosotros (he aquí una tautología; es precisamente esto lo que hace que sean inquietudes). Pero su influencia puede cambiar, y una de las cosas que pueden hacerla cambiar son las discusiones, los argumentos y la conciencia de la dirección en la que nos empujan otras inquietudes. Por lo tanto, hemos encontrado un tipo de argumento de naturaleza práctica, que toma la forma de una reflexión acerca de lo que hay que hacer, o de los principios que hay que suscribir, o de qué rasgos de carácter debemos admirar o rechazar. ¿Qué debemos pensar acerca de todo ello?

El razonamiento práctico

Al principio del capítulo hemos mencionado los razonamientos técnicos, en los cuales el objetivo es algo dado y el problema consiste en encontrar los medios para alcanzarlo. Pero está claro que buena parte del razonamiento práctico trata de influir en los objetivos de las personas. Nos esforzamos en presentar la situación bajo una luz distinta para que los demás puedan compartir los objetivos que nosotros aprobamos, o para que abandonen otros que rechazamos.

Por supuesto, este tipo de razonamientos se basa en buena medida en la persuasión. Son las artimañas del vendedor o de la agencia de publicidad. Desplegamos nuestra retórica para excitar las emociones de la gente y conducirla por los canales deseados. El predicador que pinta los horrores del infierno o el político que colorea las virtudes de su partido y los vicios del otro no están buscando mejorar el conocimiento de alguien sobre algo. Podríamos decir que

en estos casos el interés es manipular a la gente, no instruirla. El objetivo es cargar ciertas acciones de contenido emocional, para conducir a la gente en la dirección deseada. En su nivel más bajo, se trata de asociar castigos y amenazas a ciertas formas de conducta, más que de cualquier otro método más elaborado de persuasión.

Cuando adoptamos este tipo de posturas frente a los demás, lo que estamos haciendo es tratarlos como medios para nuestros fines. Por algún motivo, queremos que persigan determinado objetivo. Queremos que compren nuestro producto, que voten a nuestro partido o que acudan a nuestra iglesia. Si estamos dispuestos a adoptar cualquier procedimiento para obtener este resultado, les estamos tratando como lo que Kant llamaba «simples medios» para nuestros fines.⁶ Al manipularlos —sea mediante el engaño u otras técnicas de persuasión— confiamos en poder apartarlos de su camino, tal como confiamos en apartar cualquier otro obstáculo que se ponga entre nosotros y nuestro objetivo.

Buena parte de la vida es así, aunque no la mejor parte. Podemos adoptar posturas más cooperativas y respetuosas hacia los demás. Aunque esté convencido de que tu vida avanza a trompicones por el camino equivocado, tal vez no esté dispuesto a emplear ninguna de estas viejas técnicas para manipularte. Incluso si me dieran una vacuna mágica que te hiciera marchar por la ruta que yo deseo, a diferencia del vendedor o del predicador, yo no la aceptaría. Si hiciera eso no estaría respetando tu punto de vista y, por lo tanto, tampoco a ti como persona. Quiero que compartas mi forma de ver tu situación de la forma correcta, no por medio de la manipulación, el subterfugio, las amenazas o la fuerza bruta. ¿Cuál es esta forma correcta?

En pocas palabras: será una forma que respete y tenga en cuenta tu punto de vista. Hay cosas que claramente deben ser excluidas: el engaño y la manipulación. Hay otras que claramente hay que incluir: una buena comprensión de la situación, por ejemplo. Si yo sé cómo están las cosas y tú no, cooperaré contigo mientras busco el modo de

cambiar tu forma de actuar, siempre y cuando lleguemos a compartir la misma forma de comprender la situación.

Tal vez podríamos pensar que esto es todo, que la razón entendida como algo opuesto a la retórica debe limitarse a señalar los hechos que intervienen en la situación. Podríamos presentar un argumento como el siguiente en contra de esta conclusión. Supongamos que ambos entendemos la situación del mismo modo, que resulta ser el correcto. Supongamos entonces que yo tengo un conjunto de inquietudes que eventualmente se resuelven en cierto objetivo. ¿Cómo podrías intentar cambiarlo sin recurrir a algún procedimiento de persuasión o manipulación? Por mucho que adoptes una postura cooperativa, ¿acaso no estamos en conflicto, desde el momento en que mis inquietudes definen mi forma de reaccionar y tú deseas alejarme de alguna de ellas? No puedes hacerme cambiar respetando mis inquietudes, ya que la suposición de partida era precisamente que éstas señalaban en una dirección que no te gustaba.

Por suerte, este argumento tiene dos lagunas importantes. La primera deriva del hecho de que nuestras inquietudes no son siempre evidentes para nosotros mismos. Es posible que nuestra forma de reaccionar ante la situación no refleje de forma adecuada todo aquello que nos importa. Cuando «les damos la vuelta a las cosas» en nuestra mente, es como si explorásemos a nuestro alrededor para ver si hay algún aspecto que no hubiéramos tenido en cuenta, y que fuera importante para nosotros. Al mismo tiempo, tratamos de descubrir si interviene algún factor del que no somos conscientes: si es posible que algún aspecto tenga más importancia de la que pretendemos darle. Podemos ser tan ciegos ante nuestra propia naturaleza como lo somos a veces ante el mundo que nos rodea. Una conversación que intente descubrir motivaciones que hayamos suprimido o descartado constituye una forma de colaboración, no de manipulación.

En segundo lugar, incluso si comprendemos adecuadamente

nuestra situación y nuestras inquietudes nos resultan suficientemente transparentes para nosotros mismos, no tenemos por qué pensar que nos están manipulando o tratando de persuadirnos de algo por el simple hecho de que nos propongan otra forma de interpretar la situación. Pensemos en un caso que tenga cierto contenido moral. Por ejemplo, supongamos que optas por una determinada línea de acción que desde mi punto de vista no refleja adecuadamente el deber de gratitud o lealtad que tienes hacia una tercera persona. Te lo digo. Estoy poniendo mis cartas sobre la mesa: no hay ningún engaño o manipulación. Y puede ser que cambie tu forma de actuar, ya que si me respetas lo suficiente es posible que mi opinión influya sobre ti; por otro lado, aunque te sientas inclinado a ignorar esta opinión y seguir comportándote igual, no deja de ser un elemento más para que tengas en cuenta.

Este segundo mecanismo es, en cierto sentido, una forma de hacerte ver otro factor que interviene en la situación: mi propia desaprobación ante tu manera de actuar. Pero está claro que mi intención no es que la cosa termine ahí. Si así fuera, mi desaprobación funcionaría como un «objeto»: un mero obstáculo frente a tu causa preferida, que podrías manejar como un factor más en un análisis coste-beneficio. Pero no es eso lo que yo persigo. En la discusión moral cooperativa, lo que se pretende es llegar a un terreno común en el que plantear la aprobación y la desaprobación. He puesto mi desaprobación sobre la mesa como algo que deberías compartir o criticar, pero en cualquier caso como algo que debes considerar en los términos en que se plantea. De otro modo, mi desaprobación estaría siendo objetivada, como la inquietud de Annie por su carrera en el ejemplo que vimos antes.

Así pues, la discusión pretende establecer si debemos reconocer un deber de gratitud o lealtad o si en este caso se trata de alguna otra cosa: tal vez un fetiche que sea preferible ignorar o dejar de lado. Al tratar de responder a esta pregunta descubrimos otras cosas

que nos importan. Podríamos traer a colación, por ejemplo, consideraciones sobre lo mal o bien que iría el mundo si nadie se preocupara por estas cosas. O podríamos intentar relacionarla con otras cosas que nos importan, como la amistad o la honestidad.

Por debajo de este método discurre una inquietud fundamental: que nuestras posturas deberían ser coherentes. Y tal vez deberían ser otras cosas también, como por ejemplo imaginativas y objetivas.

Coherencia, objetividad e imaginación

Buena parte de los razonamientos prácticos tratan de establecer qué es lo que nos importa realmente. Cuando ofrecemos una razón o una justificación a otra persona, intentamos presentar nuestra acción bajo una luz favorable. Algunos autores recelan ante cualquier exigencia de orden o sistematicidad en este proceso. Pretenden negar que la vida práctica dependa de ninguna clase de «reglas» o «principios». Deberíamos concebirla más bien en términos estéticos. Podemos pronunciarnos acerca de un cuadro sin apelar a ninguna clase de principios generales y articulados que justifiquen nuestro veredicto. También deberíamos recordar en este punto el ejemplo que vimos en el capítulo 1 acerca de nuestra habilidad para reconocer cosas y nuestra capacidad de identificar que una frase es gramatical, dos cosas que parecen depender del uso de principios generales o reglas, al menos a un nivel consciente.

Sin embargo, el razonamiento práctico en general no funciona de este modo. El motivo es que necesitamos saber a qué atenernos. Sucede lo mismo que en el caso de un sistema legal. No serviría de nada tener un conjunto de leyes que no se articularan según principios y reglas generales, sino que se empeñaran en «tratar cada caso según sus circunstancias». Si no pudiéramos prever por anticipado qué cosas cuentan como méritos, no podríamos regular nuestro modo de vivir de acuerdo con este «sistema». Simplemente

no habría ley. Ocurre lo mismo con la ética. Debemos saber a qué atenernos, y esto significa identificar ciertos aspectos en las diversas situaciones o escenarios de las decisiones que apoyan o desmerecen una determinada decisión o actitud práctica. Esto significa que mientras nuestros deseos y caprichos pueden ser tan volubles como queramos, las actitudes que exigimos a las demás personas no pueden serlo. Necesitan articularse en un sistema defendible.

En el capítulo 6 vimos que la lógica valora la consistencia por encima de cualquier otra cosa. Tiene que haber una forma de que nuestras creencias sean verdaderas. En la vida práctica, la virtud equivalente consiste en que tiene que haber una forma de que se puedan implementar todos nuestros valores. Un sistema de leyes es inconsistente cuando es imposible obedecer a todas las limitaciones que impone (supongamos, por ejemplo, que prohíbe la consumición de alcohol los domingos pero que al mismo tiempo obliga a acudir a misa, en un servicio que incluye vino). Ahora bien, la vida ofrece un buen número de ejemplos en los que parece haber una inconsistencia entre valores simples. Ve siempre con la verdad por delante; no hagas daño a nadie. Pero lo cierto es que en más de una ocasión la verdad resulta dolorosa. Respeta siempre la propiedad; no pongas en peligro al Estado. Pero puede haber ocasiones en que la protección del Estado requiera confiscar la propiedad de alguien. Por lo tanto, buena parte del pensamiento práctico consiste en ajustar las obligaciones y las limitaciones que podemos exigirnos unos a otros, en resolver conflictos y complejidades y en tratar de descubrir el mejor modo de aproximarnos a una forma de vida completa y consistente. No es una tarea fácil, y sus resultados tienden a ser aproximativos y provisionales, siempre supeditados a los nuevos casos y problemas que se puedan presentar.

Por suerte, disponemos de algunas estrategias que nos pueden servir de ayuda. Una de ellas es la historia, concebida en términos de la supervivencia de los más adaptados. Las modificaciones y

soluciones que encarna nuestra forma de vida heredada tienen a su favor que han sobrevivido a la prueba del tiempo. Tenemos que prevenirnos ante la clase de adoración conservadora de las prácticas heredadas que se asocia a pensadores como Edmund Burke (1729-1797).⁷ Pero es mucho menos inteligente bascular hacia el otro extremo y creer que el test del tiempo no demuestra nada. Al fin y al cabo nos ofrece una base de datos desde la cual pensar en posibles cambios. Otra estrategia que puede resultar útil es la imaginación. No hace falta esperar la llegada de una crisis, cuando podemos acudir sin el menor coste a la ficción y a la imaginación, o a la simple actitud de plantearnos nuestros valores y la importancia que merecen en relación con otros. Y esta clase de reflexiones se plantean más a menudo cuando tenemos una perspectiva objetiva sobre la situación —cuando nos vemos a nosotros mismos tal como nos ven los demás— que en el momento en que nos empuja la pasión o cuando nos encontramos en plena acción. Esta forma de pensar nos permitiría conseguir una mejor comprensión de nuestras ideologías e hipocresías.

El relativismo

Así pues, ¿a fin de cuentas «sólo quedamos nosotros»? ¿Acaso nuestros queridos valores e imperativos morales quedan reducidos a un conjunto de inquietudes contingentes, localizadas, a lo mejor variables, que tenemos por costumbre exigir a los demás?

Pues bien, es cierto que somos nosotros, pero tal vez no «sólo» nosotros. Este «sólo» insinúa que hay otras soluciones igualmente buenas, o igualmente «válidas» o valiosas. En algunos casos particulares podríamos llegar a estar de acuerdo con esto. Los británicos conducen por la izquierda y los americanos por la derecha. Ambos países han encontrado una buena solución para el problema fundamental de la coordinación del tráfico. Conducir por uno de los

dos lados y no por el otro es algo que hacemos «sólo nosotros». Pero no somos sólo nosotros los que conducimos exclusivamente por uno de los dos lados. Conducir alternativamente por uno y otro lado, o por en medio, no es una solución tan buena —en realidad, no es ninguna solución— para el problema del tráfico.

Cuando vemos que una solución es sólo una entre muchas soluciones posibles e igualmente buenas para resolver un determinado problema, podemos pensar que es «sólo nuestra». Y ya no tendremos la tentación de moralizar a los demás acerca de ella. Los diferentes idiomas disponen de palabras diferentes para designar la misma cosa, y diferentes gramáticas y órdenes sintácticos, pero hasta ahora parece que todos han servido al propósito de la comunicación con la misma eficacia. Podemos ver las diferentes costumbres, ritos, prácticas y planificaciones sociales de todo tipo como soluciones distintas para similares problemas de expresión pública, coordinación y comunicación. No tenemos por qué establecer jerarquías. Allí donde fueres, haz lo que vieres.

Pero supongamos que una sociedad resuelve sus problemas de manera que hacen chirriar nuestras convicciones más íntimas. Supongamos que, al igual que los talibanes en el Afganistán contemporáneo, hay sociedades que niegan la educación a las mujeres. O supongamos que la historia les ha legado un sistema de castas que niega cualquier derecho a la salud, la educación e incluso el sustento a grupos enteros de gente, en función de su nacimiento. O que el paso de los siglos les haya legado un sistema en el cual ciertas personas son dueñas del cuerpo y el alma de otras. Tales sistemas ofrecen soluciones al problema de cómo vivir. Pero no tenemos por qué verlas como igualmente buenas («diferentes sin más»), o ni tan siquiera tolerables. Tenemos derecho a pensar que estas sociedades pisan los límites que nos importan. Traspasan las fronteras del respeto y de los valores que creemos que deberían protegerse. En este punto, es natural recurrir al lenguaje de los

«derechos», con el que no sólo se quiere decir que es bueno o bonito que la gente muestre preocupación y respeto por los demás, sino que en caso de que no fuera así, las partes lesionadas tienen todo el derecho de sentirse agredidas y hacer un llamamiento al mundo para que rectifique la situación.

Al decir estas cosas, damos voz a nuestras simpatías, inquietudes y valores. Pero el razonamiento práctico no puede ir más allá. No hay razón para que nos sintamos culpables por ello, como si necesitáramos algún certificado firmado por Dios o por alguna verdad normativa (lo que Platón llamaba las «formas») para tener derecho a defender nuestras opiniones. Nuestras inquietudes éticas quedan bien reflejadas en el modelo del barco de Neurath (capítulo 1). Debemos inspeccionar cada parte, y debemos hacerlo mientras confiamos en las demás partes. Pero el resultado de esta inspección bien podría ser, si somos coherentes e imaginativos, que el barco está en perfectas condiciones para navegar. E incluso si durante la travesía nos cruzamos con otros barcos que navegan en otras direcciones, no tenemos por qué lamentar que no estemos varados en algún dique seco certificado por la Razón o por Dios. Por otro lado, no existe ningún dique de este tipo.

Para terminar

En este libro he intentado presentar algunos de los grandes temas, así como aquello que podemos pensar y aquello que otras personas han pensado acerca de ellos. He procurado evitar imponer un determinado conjunto de doctrinas o puntos de vista al lector. De hecho, el lector sensible puede haber notado que los argumentos suelen desembocar en una suerte de pesimismo. La armonía entre nuestros pensamientos y el mundo, el puente que construimos entre pasado y futuro, nuestra idea de lo que contiene el mundo físico y de cómo nuestras mentes se ajustan a él, son problemas contra los que

han arremetido los pensadores más perspicaces, sin obtener éxitos convincentes. Siempre creemos adivinar cosas mejores que podríamos decir, con sólo encontrar las palabras adecuadas.

Sería posible mostrarse cínico acerca de esto —algunos filósofos profesionales son conocidos por serlo—, como si la defensa de la reflexión crítica que ofrecí en la «Introducción» se hubiera revelado vacía de contenido. No creo que esté justificada tal actitud. Creo que el proceso de enfrentarnos a los problemas es un bien en sí mismo. Si el resultado es lo que Hume llama un «escepticismo mitigado»,⁸ es decir, nos hace conscientes de lo modestos que nos conviene ser acerca de nuestras especulaciones intelectuales, acaso sea buena cosa. El mundo está lleno de ideas, y no estará de más que aprendamos a reconocer su poder, sus dificultades, sus puntos débiles y su falibilidad.

Notas

El conocimiento

1. Descartes, *Meditaciones metafísicas y otros textos*, pág. 16.
2. *Ibíd.*, pág. 20.
3. *Ibíd.*, pág. 22.
4. *Ibíd.*, pág. 24.
5. *Ibíd.*, pág. 27.
6. *Ibíd.*, pág. 30.
7. Este experimento mental fue ideado por Hilary Putnam, en *Razón, verdad e historia*, cap. 1.
8. Citado en J. P. Stern, *Lichtenberg: A Doctrine of Scattered Occasions*, pág. 270.
9. Este argumento cartesiano aparece en la tercera *Meditación*.
10. El fragmento pertenece a una carta de Descartes a Marin Mersenne, citada en *Meditations on First Philosophy*, pág. 32, nota al pie.
11. La objeción de Arnauld se halla en la tercera sección de sus *Objections and replies*, en *Descartes: Selected Philosophical Writings*, pág. 142.
12. Hume, *Investigación sobre el conocimiento humano*, Sección XII, pág. 177.
13. La metáfora de Neurath aparece en su *Anti-Spengler*.
14. Las consideraciones escépticas de Russell acerca del tiempo aparecen en *An Outline of Philosophy*, págs. 171-172.
15. La cuestión de la probabilidad y la entropía aparece en Huw Price,

Time's Arrow and Archimedes Point, cap. 2.

La mente

- 1.Las ideas de Descartes acerca del sistema nervioso aparecen especialmente en la sexta *Meditación*.
- 2.Ryle empleó esta expresión en su obra *Concept of mind* [trad. cast. *El concepto de lo mental*}. Hay que señalar que Descartes negó explícitamente que el alma estuviera ubicada en el cuerpo «como un piloto en un barco», de modo que se puede plantear la cuestión académica de si la propuesta de Descartes era más sofisticada.
- 3.Wittgenstein, *Investigaciones filosóficas*, 293, pág. 245.
- 4.Locke, *Ensayo sobre el entendimiento humano*, II. viii. 13, pág.115.
- 5.Leibniz, *Nuevos ensayos sobre el entendimiento humano*, pág. 141.
- 6.Locke, *Ensayo*, IV. x. 16, pág. 629.
- 7.Un buen manual de referencia acerca del prudente renacimiento que experimentan en la actualidad las técnicas de análisis se puede encontrar en Jackson, *From metaphysics to Ethics*.
- 8.Wittgenstein, *Investigaciones*, Pt. II. xi, pág. 475.
- 9.La mejor exposición de los avances recientes de la ciencia del color se encuentra en C.L. Hardin, *Color for Philosophers*.

La libertad

- 1.Véase libro II, 11. 251-257, pág. 74.
- 2.Schopenhauer, *On the freedom of the Will*, pág. 43.
- 3.Véase Spinoza, *Ética*, IV, V.
- 4.Véase Aristóteles, *Ética Nicomaquea*, III, 5(1114a4).
- 5.La idea de Strawson apareció por primera vez en su celebrado ensayo titulado *Libertad y resentimiento*.
- 6.Historia adaptada de Shah, *Cuentos de los Derviches*.
- 7.Wittgenstein, *Aforismos, cultura y valor*.

El yo

- 1.Hume, *Tratado de la naturaleza humana*, I. iv. 6, pág. 355.
- 2.Reid, *Essays on the Intellectual Powers of Man*, pág. 202.
- 3.Locke, *Ensayo*, II. xxvii. 4, pág. 313.
- 4.Locke, *Ensayo*, II. xxvii. 12, pág. 320.
- 5.Kant, *Crítica de la razón pura*, A364, pág. 341.
- 6.Reid, *Essays*, pág. 213
- 7.Locke, *Ensayo*, II. xxvii. 20, 326.
- 8.Wittgenstein, *El cuaderno azul*, pág. 104.
- 9.Ésta es la gran maniobra de Kant. Los pasajes cruciales de la *Crítica de la razón pura* se encuentran en la sección titulada «Deducción trascendental de los conceptos puros del entendimiento», B 130-B170.
10. Véase la nota anterior.

Dios

- 1.San Anselmo, *Proslogion*.
- 2.Hume, *Diálogos sobre religión natural*, P. 9, págs. 101-102.
- 3.Ibíd.,P. 9, págs. 103-104.
- 4.Ibíd.,P. 2, pág. 29.
- 5.Ibíd.,P. 2, pág. 37.
- 6.Ibíd.,P. 7, pág. 83.
- 7.Ibíd.,P. 7, págs. 88-89.
- 8.Ibíd.,P. 5, págs. 71-72.
- 9.Ibíd.,P. 10, pág. 63.
10. 10. Ibíd.,P. 11, pág. 141.
- 11.Wittgenstein, *Investigaciones filosóficas*, 304, pág. 249
- 12.Hume, *Diálogos*, P. 12, pág. 168.

- 13.Hume, *ibíd.*, X, Pt. 1, pág. 137.
- 14.Hume, *ibíd.*, X, Pt. 1, pág. 137.
- 15.Hume, *ibíd.*, X, Pt. 1, pág. 140.
- 16.Hume, *ibíd.*, X, Pt. 2, pág. 152.
- 17.Hume, *ibíd.*, X, Pt. 2, pág. 142.
- 18.Hume, *ibíd.*, X, Pt. 2, pág. 146.
- 19.Véase Pascal, *Pensamientos*, págs. 126-130.
- 20.Mill, *Sobre la verdad*.
- 21.Clifford hace tal comparación en «The Ethics of Belief», en sus *Lectures and Essays*, pág. 346.
- 22.Coleridge, *Aids to Reflection*, Aforismo XV, pág. 107.

Razonar

- 1.Las ideas presentadas han sido objeto de un estudio detallado por Paul Grice. Sus artículos están reunidos en *Studies in the Way of Words*.
- 2.Hume, *Investigación*, IV, Pt. 2, págs. 56-57.
- 3.*Ibíd.*, IV, Pt. 2, pág. 60.
- 4.El fragmento corresponde a una carta de Newton a Bentley. Aparece citada en Kemp Smith, *The Philosophy of David Hume*, pág. 61.
- 5.Hume, *Investigación*, IV, Pt. 1, págs. 49-50.
- 6.*Ibíd.*, IV, Pt. I, pág. 51.
- 7.*Ibíd.*, IV, Pt. I, pág. 53.
- 8.La obra fundamental de Kuhn es *La estructura de las revoluciones científicas*, publicada en 1962.

El mundo

- 1.Galileo, «El ensayador», en *Galileo*, pág. 113.
- 2.Descartes, *Meditaciones*, 6, pág. 76.

3. Russell plantea brevemente su acusación de «favoritismo» en *The problems of Philosophy*, cap. 1, pág. 10 (trad. cast.: *Los problemas de la filosofía*), Cerdanyola, Labor, 1991).
4. Locke, *Ensayo*, II. viii. 9 y 10, pág. 113.
5. Ibíd., II. viii. 14, pág. 115.
6. Ibíd., II. iv. I, pág. 101.
7. Ibíd., II. iv. 6, pág. 105.
8. Hume, *Tratado*, I. iv. 4, pág. 326.
9. Ibíd. I. iv. 4, pág. 327.
10. Faraday, «A Speculation touching Electrical Conduction and the Nature of Matter», debo la cita a Langton, *Kantian Humility*, pág. 101.
11. Hume, *Investigación*, VII Pt. 1, pág. 88.
12. Kant, *Crítica de la razón pura*, A368, pág. 344.
13. Ibíd. A157/B196, pág. 195.
14. Ibíd. Prólogo a la segunda edición, pág. 20.
15. Ibíd., A29, pág. 72.
16. Ibíd. A372, pág. 346.
17. Ibíd. B275, pág. 247.
18. Berkeley, *Tres diálogos*, Diálogo 1, pág. 398, págs. 134-135.
19. Este argumento aparece en sus *Principia Ethica*, cap. 6, *passim*.
20. Véase Wittgenstein, *Investigaciones filosóficas*, § 137-203, (aproximadamente, ya que se mezcla con otras cuestiones).

¿Qué hacer?

1. Resulta difícil encontrar un defensor del egoísmo psicológico puro, aunque a veces se señala a Thomas Hobbes como uno de ellos. La exposición clásica la ofrece Joseph Butler en sus *Fifteen Sermons Preached at the Rolls Chapel*, en 1726, en especial el Sermón XI. Dedico más espacio a este tema en mi libro *Ruling Passions*, caps. 5 y 6.

- 2.Hume, *Tratado*, III. iii. I, pág. 782.
- 3.San Agustín, *The Literal Meaning of the Genesis*, Libro 4, cap. 4, § 8.
La traducción ha sido levemente alterada.
- 4.El argumento de la «cuestión abierta» de Moore procedede los
Principia Ethica, págs. 10-20.
- 5.Platón celebra estos atisbos de un orden ideal y normativo como intuiciones de las «formas». Sin embargo, existe un importante debate académico acerca de lo que significa esta doctrina y hasta qué punto mantiene Platón sus posiciones entre un diálogo y otro.
- 6.Es difícil ubicar la crítica de Kant contra el trato de los demás como «simples medios» en algún pasaje concreto de su *Fundamentación de la Metafísica de las Costumbres*.
- 7.El conservadurismo de Burke se refleja en sus *Reflexiones sobre la Revolución Francesa*.
- 8.Hume se refiere en términos favorables al escepticismo mitigado en su *Investigación*, XII.

Bibliografía

- Anselmo, san, *Proslogion*, Buenos Aires, Aguilar, 1970.
- Aristóteles, *Ética Nicomáquea*, Madrid, Gredos, 1988 [trad. Julio Pallí Bonet].
- Berkeley, George, *Tres diálogos ente Hilas y Filonús*, Barcelona, Orbis, 1980.
- Burke, Edmund, *Reflexiones sobre la Revolución Francesa*, Madrid, Centro de Estudios Constitucionales, 1978.
- Descartes, *Meditaciones metafísicas y otros textos*, Madrid, Gredos, 1987.
- Galileo, Galileo. *Antología*, Barcelona, Península, 1991.
- Hume, David, *Diálogos sobre religión natural*, México, FCE, 1942.
- Hume, David, *Investigación sobre el conocimiento humano*, Madrid, Alianza, 1980.
- Hume, David, *Tratado de la naturaleza humana*, Madrid, Tecnos, 1988.
- Kant, Immanuel, *Crítica de la razón pura*, Madrid, Alfaguara, 1978.
- Kant, Immanuel, *Fundamentación de la metafísica de las costumbres*, Madrid, Espasa-Calpe, 1981.
- Kuhn, Thomas, *La estructura de las revoluciones científicas*, Madrid, FCE, 1981.
- Leibniz, G.W., *Nuevos ensayos sobre el entendimiento humano*, Madrid, Alianza, 1992.
- Locke, John, *Ensayo sobre el entendimiento humano*, México, FCE, 1956.

- Lucrecio, *De la naturaleza*, Madrid, CSIC, 1983.
- Mill, John Stuart, *Sobre la libertad*, Madrid, Alianza, 1970.
- Moore, George Edward, *Principia Ethica*, México, Universidad Nacional Autónoma, 1959.
- Pascal, Blaise, *Pensamientos*, Madrid, Alianza, 1981.
- Putnam, Hilary, *Razón, verdad e historia*, Madrid, Tecnos, 1988.
- Russell, Bertrand, *Los problemas de la filosofía*, Barcelona, Labor, 1952.
- Ryle, Gilbert, *El concepto de lo mental*, Barcelona, Paidós, 1967.
- Schopenhauer, Arthur, *Sobre la libertad humana*, Madrid, Revista de Occidente, 1934.
- Shah, Idries, *Cuentos de los Derviches*, Barcelona, Paidós, 1994.
- Spinoza, *Ética*, Madrid, Alianza, 1987 [trad. Vidal Peña].
- Strawson, Peter E, *Libertad y resentimiento*, Barcelona, Paidós, 1995 [trad. Juan José Acero].
- Wittgenstein, Ludwig, *Aforismos, cultura y valor*, Madrid, Espasa-Calpe, 1995.
- Wittgenstein, Ludwig, *Investigaciones filosóficas*, Barcelona, Crítica, 1988.
- Wittgenstein, Ludwig, *Los cuadernos azul y marrón*, Madrid, Tecnos, 1968.