

Mackie

## V

### El objeto de la ética

#### 1. Consecuencias del escepticismo moral

He argumentado en el capítulo 1 que no existen valores objetivos, y en los capítulos 2, 3 y 4 he afirmado que no pueden derivarse conclusiones morales sustantivas ni hacerse restricciones serias sobre los puntos de vista morales tanto si se intenta hacerlo partiendo de los significados de los términos morales como si se intenta a partir de la lógica del discurso moral. ¿Qué tareas quedan entonces para la filosofía moral? Uno podría estudiar los puntos de vista morales así como las creencias de nuestra propia sociedad o de otras, procediendo quizás a un estudio diacrónico y haciendo nuestro el tema que se resume en el título de Westermarck, *Origen y desarrollo de las ideas morales*. Sin embargo, esto quizá pertenece más bien a la antropología o a la sociología. Más afín a los filósofos y más adaptable a los métodos filosóficos sería el intento de describir sistemáticamente nuestra conciencia moral o una parte de ella, tal como nuestro «sentido de la justicia», el intento de encontrar algún conjunto de principios que fuera en sí mismo razonablemente aceptable para nosotros y con el cual, añadido a sus consecuencias prácticas y a sus aplicaciones, nuestros detallados e «intuitivos» (aunque subjetivos en realidad) juicios morales pudieran hallarse en «equilibrio reflexivo». Es decir, debemos empezar al mismo

tiempo con unos principios generales que sean aceptables *prima facie* y con la masa de juicios morales detallados también aceptables *prima facie* con la intención de ajustarlos, allí donde no concuerden por completo, hasta que se alcance el más satisfactorio y coherente de los compromisos. Esto es lo que John Rawls llama «una teoría de la justicia» (en el libro así titulado). Se trata de una investigación legítima, pero no debe confundirse con el intento, superficialmente similar pero fundamentalmente distinto en su propósito, de pensadores como Sidgwick que han tratado de abrirse paso mediante nuestras diversas «intuiciones» hasta una verdad moral objetiva, hasta una ciencia de la conducta. «Nuestro sentido de la justicia», tanto si se trata del suyo como si se trata del mío, y también si se trata del de un grupo mucho más amplio, no tiene autoridad sobre aquellos que disienten de sus recomendaciones y ni siquiera la tiene sobre nosotros en el caso de que sintamos inclinación a cambiar de parecer. Pero si no existen verdades morales objetivas que descubrir, ¿ya no queda otra cosa que hacer excepto describir nuestro sentido de la justicia?

Al menos podemos considerar el asunto desde otro ángulo. La moral no es algo que deba descubrirse sino algo que debe hacerse: tenemos que decidir cuáles serán los puntos de vista morales que adoptaremos, las normas morales que habremos de seguir. No hay duda de que las conclusiones a las que lleguemos reflejarán y revelarán nuestro sentido de la justicia, nuestra conciencia moral, es decir, el estado de nuestra conciencia moral al final de la discusión, no necesariamente el aspecto que tenía al comienzo. Pero no es éste el objeto del ejercicio: el objeto es más bien decidir qué hacer, qué defender y qué condenar, qué principios de conducta aceptar y apadrinar en tanto guías o formas de controlar nuestras propias elecciones y quizá también las de otras personas.

Con todo, incluso en el caso de que estemos considerando de este modo la moral, hay una distinción que debe hacerse. Una moral, en sentido lato, será una teoría de la conducta general, incluyente: la moral que alguien respalda será cualquier cuerpo de principios a los que esa persona permita en último término guiar o determinar sus opciones de acción. En sentido estricto, una moral es un sistema que define un particular tipo de restricciones sobre la conducta, unas restricciones cuya tarea principal es proteger los intereses de personas distintas al

propio agente y que se presentan ante el agente como desafíos para sus inclinaciones naturales o para sus tendencias espontáneas de acción. En este sentido estricto, las consideraciones morales serán consideraciones de una cierta gama restringida y no incluirán necesariamente todo lo que un hombre permite que determine sus actos. En este segundo sentido, alguien podría decir deliberadamente: «Admito que la moral requiere que yo haga esto y aquello, pero no tengo intención de hacerlo: para mí existen aquí otras consideraciones que desbordan la consideración moral». Y no necesita poner la palabra «moral» entre comillas, ya sea tácita o explícitamente. Muy bien podría estar hablando de su propia moral, de las restricciones morales que él mismo acepta y respalda como tales por lo general. Pero dado que en este sentido estricto las consideraciones morales son sólo unas consideraciones entre otras a las que también presta su respaldo y no un sistema incluyente que incorpore y, allí donde sea necesario, sopesé comparativamente todas las razones que este hombre acepta como razones a favor o en contra de hacer algo, es posible que en cierta situación particular las consideraciones morales deban verse desbordadas. Pero nadie puede, en lo que se refiere a sus opciones de acción, desbordar deliberadamente lo que se halla en su moral en sentido amplio, pese a que pudiera apartarse de ella por causa de una «voluntad débil».

No tiene interés discutir cuál es el sentido más correcto de «moral», si el sentido estricto o el sentido amplio. Los dos se utilizan y los dos manifiestan tener raíces y vínculos importantes en nuestro pensamiento. Pero es esencial no confundirlos, no pensar que lo que reconocemos como consideraciones (en el sentido estricto) peculiarmente morales poseen, de manera necesaria (saltando así al sentido amplio) una autoridad última sobre nuestras acciones. No debemos suponer que cualquier sistema general de principios de elección que podamos aceptar mediante adecuada reflexión deba construirse completamente a partir de materiales que podríamos llamar morales en el sentido restringido.

Sin embargo, estoy seguro de que algo del tipo de lo que acabo de definir (sólo de forma muy aproximada) como moral en sentido estricto constituirá una parte importante de cualquier moral en sentido amplio que sea aceptable mediante reflexión. Quiero por ello examinarla más de cerca para ver en qué con-

siste su naturaleza y qué es lo que cuenta de hecho en la tendencia que tiene a usurpar tanto el nombre como la función de una teoría general de la conducta.

## 2. Un instrumento para contrarrestar las simpatías limitadas

Si preguntamos para qué sirve o cuál es el objeto de la moral, G. J. Warnock cree que lo entenderemos mejor como moral en sentido estricto. La moral es una especie de valoración, un tipo de estimación de la conducta humana. Esto debe, según sugiere Warnock, poseer algún rasgo distintivo, debe haber algo que supuestamente esté haciendo aflorar. Warnock lo explica en los términos de unos ciertos rasgos generales y persistentes que se dan en los apuros que padecen los hombres y que se resumen en que «inherentemente, las cosas podrían ir realmente mal», mal en el sentido natural, no en el sentido moral, es decir, en el sentido de que los deseos humanos, las necesidades y los intereses tienen grandes probabilidades de verse frustrados en buena medida. Entre los factores que contribuyen a hacer que las cosas vayan mal en cuanto al curso natural de los acontecimientos se encuentran varias limitaciones: recursos limitados, información limitada, inteligencia limitada, racionalidad limitada..., pero por encima de todo, simpatías limitadas. Los hombres se manifiestan a veces una mutua malevolencia activa, pero incluso dejando aparte este hecho, casi siempre se preocupan más de sus fines egoístas que de ayudarse unos a otros. La función de la moral es primordialmente la de contrarrestar esta limitación de las simpatías humanas. Podemos decidir que el contenido de la moral debe hallarse en la investigación de cuál es el mejor modo de realizar esta función.

Éste es un enfoque útil, que ha sido puesto de manifiesto por numerosos pensadores. Existe una versión colorista en el diálogo platónico de *Protágoras*, en donde el sofista Protágoras incluye esta idea en una explicación deliberadamente mítica de la creación y los albores de la historia humana. En el momento de su creación, los hombres estaban, en comparación con los de-

más animales, pobremente equipados. No tenían las garras ni la fuerza ni la velocidad ni la gruesa piel o las escamas de otros animales para poder encontrar comida y protegerse de los enemigos y los elementos. Para compensar esto, se les concedió el uso de las diversas artes mecánicas y la utilización del fuego, de modo que pudiesen hacer casas, vestidos, herramientas y armas y fuesen capaces de producir comida por sí mismos. Sin embargo, a pesar de esto, se hallaban en una situación francamente mala, ya que cuando vivían dispersos en pequeños grupos no eran rivales para las bestias salvajes y cuando se reunían se maltrataban unos a otros debido a que carecían «del arte político». Al final, Zeus se apiadó de ellos y ordenó a Hermes que entregara a los hombres *aidós* (que quizá pudiéramos traducir por «sentido moral») y *dike* (la ley y la justicia) para que se convirtieran en los principios ordenadores de las ciudades y en la base para los lazos de la amistad. Separada de su marco mitológico, la tesis de Protágoras es clara: es preciso disponer de un sentido moral, de la ley y la justicia, para permitir que los hombres vivan juntos en comunidades suficientemente amplias para competir con éxito con los animales salvajes.

Hobbes traza un cuadro similar, excepto en el hecho de que olvida las bestias y de que el peligro viene ahora del daño que los hombres puedan inferirse mutuamente. La competencia y la desconfianza mutua hacen que el estado natural del hombre sea una guerra de todos contra todos. Pero Hobbes es mucho más explícito que Protágoras con respecto a la solución. El miedo a la muerte y el deseo tanto de autopreservarse como de adquirir, por ingenio, los medios para una vida decente y para asegurar lo que se posee llevan a los hombres a buscar la paz. Pero nadie puede dejar de luchar contra los demás a menos que también los demás dejen de luchar contra él. De este modo, lo que se necesita para la paz es un acuerdo para poner límites a las exigencias antagónicas. Sin embargo, incluso si se efectúa un acuerdo semejante, nadie tiene motivo suficiente para atenerse a él a menos de que tenga la necesaria seguridad de que las otras partes también lo acatarán. De ahí que el único tipo de acuerdo que logrará este propósito de establecer la paz (y de hacer con ello más segura y confortable la vida) es aquel que instaure un mecanismo para respaldar el propio acuerdo. Este mecanismo, según Hobbes, debe estar encarnado en un soberano político, un hombre (o una corporación) que

por sí mismo no sea una de las partes del acuerdo pero al cual todas las partes consensuen obediencia. Los principios morales que Hobbes ofrece como solución necesaria para el problema de la competencia y la desconfianza natural se contienen en una serie de leyes de la naturaleza: procurar la paz si existe la esperanza de alcanzarla, aceptar la mutua limitación de las exigencias antagónicas, llegar a acuerdos, mostrar gratitud a cambio de los beneficios, adaptarse cada uno a los demás, perdonar las ofensas pasadas de aquellos que se arrepientan y den seguridad de no repetir sus ofensas, evitar el castigo retrospectivo y de carácter retributivo, y ocho «leyes» más del mismo tipo general. Sin embargo, esto solo no es suficiente. El mecanismo esencial es una forma de acuerdo que proporciona su propio refuerzo. Cada una de las partes tiene un motivo para soportar la autoridad de quien tiene el encargo de castigar las brechas abiertas en el acuerdo (y que, a su vez, tendrá, él mismo, un motivo para hacerlo). Por consiguiente, cada una de las partes tendrá una doble razón para cumplir su participación en el trato: el miedo al castigo por incumplirlo y la expectativa de beneficios si lo mantiene, ya que el cumplimiento de las otras partes de sus porciones del trato queda suficientemente bien asegurado por los mismos motivos. Aunque Hobbes habla de cómo los hombres en estado de naturaleza se reúnen y, mediante dicho contrato social, establecen una sociedad civil y un poder soberano, podemos considerar este contrato históricamente original como algo de entidad no menos mítica que la intervención de Zeus y Hermes en la narración de Protágoras. Sin embargo, podemos aceptar o al menos considerar seriamente todo lo esencial de la explicación de Hobbes. Puede sostenerse que semejante tipo de contrato está implícito en las sociedades humanas, y que lo está necesariamente, ya que la decadencia de las relaciones y motivaciones sobre las que Hobbes llama la atención podría conducir a profundos conflictos y a la radical inseguridad de la vida.

Hume representa otra tradición: «Sólo del egoísmo y la limitada generosidad del hombre, añadido a la escasa provisión que la naturaleza ha generado para la satisfacción de sus deseos, deriva su origen la justicia». La justicia (término con el que Hume indica particularmente el respeto a la propiedad y a las reglas para administrar su posesión y su transferencia, es decir la honestidad y el mantenimiento de las promesas) es una virtud artificial. No es algo que nos sintamos naturalmente

te inclinados a aprobar, sino un mecanismo que es beneficioso debido a ciertos rasgos contingentes de la condición humana. Si los hombres hubieran resultado ser abrumadoramente benevolentes, si cada cual hubiera procurado únicamente la felicidad de todos, si los hombres hubiesen amado a su prójimo como a sí mismos, no hubiera habido necesidad de las reglas que conforman la justicia. Tampoco hubiera habido ninguna necesidad de ellas si la naturaleza hubiera ofrecido abundantemente y sin esfuerzo alguno por nuestra parte todo lo que hubiésemos podido desear, si el alimento y el refugio hubiesen sido cosas inagotablemente disponibles, tal como, hasta hace bien poco, parecían ser el aire y el agua. La realización y mantenimiento de promesas y tratos es un dispositivo que permite una cooperación mutuamente beneficiosa entre personas cuyos motivos son principalmente egoístas y entre las cuales es preciso que las contribuciones de las diferentes partes se hagan en momentos distintos. «Tu maíz está maduro hoy, él mío lo estará mañana. Es beneficioso para ambos que yo recoja hoy contigo la cosecha y que tú me ayudes mañana. No se trata de amabilidad hacia ti y yo sé que tampoco tú la sientes por mí. No asumiré por tanto ningún esfuerzo en tu favor, de modo que trabajaré contigo por mi propia cuenta a cambio de una ayuda similar. Sé que me frustraría y que sería en vano si esperara tu gratitud. En este punto, por tanto, te dejo trabajar solo: tú me tratarás de la misma manera. La estación avanza y tanto tú como yo hemos perdido nuestras cosechas por falta de mutua confianza y seguridad.» El mecanismo de la promesa existe con el fin de superar callejones sin salida como éste. Sin embargo, insiste Hume, un solo acto de justicia, considerado en sí mismo, puede hacer mayor mal que bien: «Es únicamente la competencia de la humanidad, según un esquema o sistema general de acción, lo que tiene sus ventajas».

### 3. La forma del instrumento

Protágoras, Hobbes, Hume y Warnock están todos más o menos de acuerdo con respecto a cuál es el problema que la moral debe resolver: la unión de recursos limitados y simpa-

tías limitadas genera al mismo tiempo la competencia que conduce al conflicto y la ausencia de lo que podría ser una cooperación mutuamente beneficiosa. Sin embargo, existen algunas diferencias en sus esquemas de solución, diferencias que son instructivas y requieren mayor profundización.

Hobbes habla de las leyes de la naturaleza, es decir, de las reglas morales básicas que son inmodificables porque las líneas maestras de los apuros humanos no cambian. Sin embargo, aduce, no son incondicionalmente válidas en tanto reglas de conducta. Todas ellas nos «imponen el deseo de que se vean cumplidas: Pero [...] en lo que hace a llevarlas a la práctica, no siempre es así». Uno no puede permitirse obedecer estas reglas a menos que tenga alguna garantía de que los demás también lo harán; pero si uno dispone de semejante garantía, entonces está obligado a obedecerlas, dado que obedecerlas representa entonces la mejor oportunidad de preservar la propia vida. Por lo tanto, desde el punto de vista de Hobbes, esas reglas, que funcionan únicamente como reglas morales, no son suficientes. Deben verse complementadas por el dispositivo político de la soberanía. Sólo si el agente sabe que esas reglas serán reforzadas, que las violaciones que otros puedan infligirles en su perjuicio serán objeto de disuasión mediante una amenaza o castigo eficaces, tendrá una buena razón para obedecerlas él mismo. Hume también habla de tres leyes fundamentales de la naturaleza, «la de la estabilidad de la propiedad, la de su transferencia por consentimiento, y la del cumplimiento de las promesas». «La paz y la seguridad de la entera sociedad humana, dice, dependen de la estricta observancia de estas tres leyes de la naturaleza [...]». La sociedad es absolutamente necesaria para el bienestar de los seres humanos y, a su vez, éstos son igual de necesarios para el mantenimiento de la sociedad». También él ve que el respaldo que la institución de gobierno, los magistrados y sus oficiales brindan a las leyes es un imprescindible mecanismo complementario, pero tiene menos claro, o es menos ingenuo que Hobbes en cuanto a por qué resulta necesario. Hume escribe como si obedecer las reglas que hacen posible la sociedad y la cooperación constituyera un interés de cada cual a largo plazo, y como si sólo fuera la debilidad humana de preferir intereses menores inmediatos a los mayores pero más lejanos lo que se necesitara contrarrestar mediante la creación de un grupo especial de personas cuyo in-

terés inmediato fuera reforzar lo que constituye el interés de todos a largo plazo. No señala en cambio, como hace Hobbes, que aunque la conformidad general hacia estas reglas beneficie a todos, la conformidad que unilateralmente brinda no corre pareja a mi beneficio a largo plazo: por muy lejos que sean, siempre necesitarán un dispositivo de refuerzo que proporcione a cada hombre un motivo egoísta para obedecer esas reglas morales. Sin embargo, Hume también plantea la cuestión de «por qué [y cómo] unimos la idea de virtud a la de justicia, y la de vicio a la de injusticia», lo que en su teoría de la virtud y el vicio significa «por qué y cómo sentimos aprobación hacia las acciones justas y la desaprobación hacia las injustas, incluso en el caso de que no tengan efecto sobre nuestros intereses». La virtud y el vicio son objetivaciones de una aprobación o desaprobación desinteresada. Hume sugiere que los tipos de cálculos interesados ya mencionados inducen a los hombres a «ponerse a sí mismos bajo el mandato de esas reglas, ya que pueden hacer que el comercio sea más seguro y más cómodo». Pero ponerse a sí mismos bajo el mandato de reglas no es asunto tan sencillo para un grupo de hombres egoístas. Hume sugiere entonces que la simpatía, la tendencia natural a compartir los sentimientos de los directamente afectados, nos hace ampliar lo que inicialmente era una aprobación o desaprobación interesada hacia casos alejados de nuestros intereses. Esta tendencia natural se estimula posteriormente mediante, por un lado, «el artificio de los políticos», que, viendo que esos sentimientos morales son beneficiosos, los fomentan deliberadamente, y por otro, por medio de la educación privada, debido a que los padres, comprendiendo que un hombre es «tanto más útil, para sí mismo y para otros, cuanto mayor sea el grado de probidad y honradez con que se le dote», inculcan estos sentimientos en sus hijos.

Aunque estos últimos argumentos son correctos, la explicación en su conjunto es inadecuada. Sus dos fundamentos —los hombres están interesados en ponerse a sí mismos bajo la acción de reglas, y existe una simpatía natural— son demasiado débiles. Sin embargo, Hume señala que las convenciones pueden crecer gradualmente conforme los hombres vayan experimentando repetidas veces las ventajas de ajustarse a ellas y las desventajas de infringirlas. Y nosotros podemos desarrollar esta pista. Por las razones mencionadas, los sentimientos



morales que «men la idea de virtud a la de justicia» permiten que florezcan los grupos sociales en los que arraigan. Por consiguiente, las presiones habituales de la evolución, las diferencias de supervivencia de los grupos en los que estos sentimientos son fuertes, ya provengan de tendencias psicológicas heredadas o de tradiciones mantenidas por la sociedad, nos ayudarán a explicar por qué esos sentimientos se han convertido en algo vigoroso y extendido. Dado que la evolución por causa de la selección natural es el recambio habitual moderno de la divina providencia, podemos llegar a esta explicación mediante una reinterpretación casi mecánica del mito de Protagoras.

Así entendida, la solución de Hume añade algo importante a la de Hobbes. Aunque, como he dicho, las leyes de la naturaleza señaladas por Hobbes pueden comprenderse como reglas morales, el motivo que impone (siempre) «el deseo de que se vean cumplidas» y el que anima (siempre) que haya un adecuado dispositivo reforzante) «a llevarlas a la práctica» no deja de ser un motivo egoísta. Hobbes no permite el desarrollo de lo que podríamos llamar instinto secundario en favor de la moral. Sin embargo, habiendo visto cuál es la benéfica función que realiza la moral, podemos entender sin dificultad cómo es posible que existan, como de hecho es claro que existen, tales instintos secundarios o tales tradiciones sociales persistentes. Si la explicación de Hobbes se complementa con el reconocimiento de este elemento secundario, la estructura social que describe pierde parte de su vulnerabilidad como castillo de naipes, edificio que de hecho se sostiene, pero en forma precaria, estando como está cada parte sostenida por otra que, a su vez, se mantiene sólo por el apoyo que otra cosa le brinda.

Si tanto Hobbes como Hume hablan de leyes de la naturaleza, es decir de reglas, Warnock argumenta que la moral, cuya tarea es la de contrarrestar las simpatías limitadas, no puede ser cuestión de reglas, es decir, asunto de requisitos rígidos o absolutos: en el pensamiento moral nos vemos obligados a sopesar los motivos, no nos limitamos simplemente a seguir reglas. Más bien, argumenta Warnock, el contenido central de la moral debe ser un conjunto de virtudes morales. Para que las cosas no salgan tan mal como sería de esperar según su inherente condición se requieren cuatro tipos de cosas: conocimiento, organización, coerción y buenas disposiciones. War-

nock concede a Hobbes la necesidad de coerción, pero añade, creo que acertadamente, que dicha coerción no será suficiente sin las buenas disposiciones, y por tanto concede a éstas cierta prioridad. Sin embargo, entre las buenas disposiciones, Warnock distingue las virtudes no morales: el carácter indutroso, el valor y el autocontrol, que simplemente contrarrestan varios tipos de debilidad humana, de las virtudes característicamente morales: no maleficencia, justicia, beneficencia y no engañar. Son estas últimas las que constituyen el núcleo de la moral.

Ciertamente, podemos tomar estas virtudes morales, una subclase de las buenas disposiciones, como algo que posee el mismo tipo de valor y es por tanto acreedor del mismo tipo de explicación evolucionista que el utilizado para los sentimientos morales. Por otro lado, no parece existir ninguna buena razón para excluir de la moral a reglas tales como las que Hobbes y Hume consignan. Podemos posponer (hasta el capítulo 7) la cuestión de si queremos que cualquiera de ellas sea una regla absolutamente rígida e inflexible. Warnock también destaca la parte que corresponde señalar a las obligaciones. Este autor considera la obligación de mantener las promesas como un caso especial de la virtud de no engañar. Sin embargo, seguramente la explicación de Hume sobre el particular sea más correcta: prometer es un dispositivo que permite a personas cuyos motivos son principalmente egoístas ayudarse mutuamente y obtener por tanto beneficios para todos, y, en general, prometer permite que una persona confíe en las acciones futuras de otra. De nuevo, es necesario algo análogo a la explicación de Hume sobre la institución de prometer para poder explicar otras obligaciones no contratuales que surgen de ciertas relaciones especiales, como las de los padres con los hijos y las de los hijos con los padres. De hecho, Protágoras tenía razón en lo principal, aunque no pareciera tenerla explícita y suficientemente en cuanto a los detalles, es decir, tenía razón al resumir la moral que los dioses otorgan a los hombres bajo los epígrafes de *aidós* y *dike*. Porque *aidós* puede abarcar los sentimientos morales, las disposiciones y el respeto hacia varias obligaciones, mientras que *dike* puede cubrir las reglas de tipo más formal que asociamos con los mecanismos político-legales para el reforzamiento de la ley y la realización del derecho positivo. Tanto uno

como otro son partes esenciales y complementarias del mecanismo de la moral.

#### 4. Análisis de la teoría de juegos

Un análisis más abstracto realizado a partir de modelos extraídos de la teoría de juegos puede confirmar las tesis que hemos reunido a lo largo de todas estas explicaciones. El punto de partida natural es lo que se conoce como dilema del prisionero. Sin embargo su significación y conexiones se obtienen mejor mediante una de las muchas variantes de la situación. Dos soldados, Tom y Dan, luchan en dos puestos de guardia contiguos intentando detener el avance enemigo. Si los dos se mantienen en sus puestos, tendrán una razonable posibilidad de mantener al enemigo a raya hasta que lleguen refuerzos y de ese modo ambos sobrevivirían. Si los dos desertan, el enemigo abrirá brecha inmediatamente, y las posibilidades que cada uno tendrá de sobrevivir serán notablemente menores. Sin embargo, si uno se mantiene en su puesto mientras el otro deserta, el desertor tendrá una posibilidad de supervivencia aún mayor que la que tendría si ambos se mantuviesen en sus puestos, mientras que el que se queda para combatir tendrá menos oportunidades de sobrevivir que las que tendría si ambos desertaran. Supongamos que ambos hombres conocen estos hechos y que cada uno hace sus cálculos de forma completamente racional con la vista puesta en su propia supervivencia. He aquí el razonamiento de Tom: si Dan se queda en su puesto, yo tendré mayores probabilidades de sobrevivir si deserto que si me quedo en mi puesto; pero si Dan también deserta, yo tendré mejores oportunidades si deserto que si me quedo; por consiguiente, sea lo que sea lo que Dan haga, haré bien en desertar. Dado que la situación es simétrica, el razonamiento de Dan es, *mutatis mutandis*, exactamente el mismo. La consecuencia es que ambos desertarán. Y sin embargo, cada uno de ellos hubiera tenido una mejor oportunidad de sobrevivir, esto es, una mejor oportunidad de lograr el objetivo que, por hipótesis, se había propuesto, si uno y otro hubiesen permanecido en sus puestos.

A primera vista esto puede parecer una paradoja. Pero de hecho, no tiene nada de paradójico. ¿Por qué deberíamos sorprendernos que dos hombres que proceden a realizar opciones de acción separadas y descoordinadas, y que procuran lograr ambos, por muy racionalmente que lo hagan, objetivos privados e inconexos, sean incapaces de lograrlos, o por qué debería sorprendernos que hubiesen tenido, cada uno, una mejor oportunidad de haber alcanzado su objetivo particular con sólo haber sido capaces de coordinar sus opciones de acción? ¿Y cómo podrían haberlo hecho? El mejor resultado para cada uno es que el otro se vea inducido a quedarse de algún modo mientras él deserta. Pero debido a que esto plantea una situación asimétrica, sólo podrá lograrse mediante alguna forma de arbitrio o de explotación. El mejor resultado simétrico, y por tanto el mejor que podría lograr cualquier mecanismo que ambos aceptasen libremente, es que ambos se queden. Pero ¿cómo puede lograrse esto? Lo que necesitan es algo que, literal o metafóricamente, ligue sus acciones. Si ambos saben que las únicas opciones abiertas son las de que ambos se queden o que ambos deserten, entonces, calculando tan racional y también tan egoístamente como antes, ambos se quedarían. ¿Pero qué es lo que serviría para ligar sus acciones? Supongamos que hacen un trato: cada uno dice «Yo me quedaré en mi puesto si tú te quedas en el tuyo». Sin embargo, si siguen siendo egoístas racionales, cada uno tendrá idénticos motivos para romper este acuerdo que ya le ataba desde el principio y desertar antes de que haya ningún acuerdo. Si uno y otro pueden verse, y ambos saben que el otro percibirá en qué momento abandona su puesto, ¿aumentará esto las probabilidades de que mantengan su acuerdo? Puede aumentarlas si Tom cree que es más probable que Dan rompa su acuerdo si ve que Tom lo rompe, y que es más probable que lo mantenga si sabe que Tom lo mantiene y viceversa. Con todo, de acuerdo con los supuestos que hemos introducido hasta el momento, ninguno de los dos tiene ningún motivo para creer esto.

Por otro lado, sería eficaz, y también racional en la situación supuesta, que ambos hombres llegaran al acuerdo de encadenarse literalmente a sus puestos de guardia. Sería racional que cada uno aceptase esta pérdida de su propia libertad de elección con tal de que la libertad de su camarada se viera limitada en igual medida. De eficacia casi igual sería algún

tipo de disciplina externa, si cada uno supiera que cualquier hombre que desertase se haría reo de un severo castigo. Pero lo importante con respecto a nuestros propósitos es que pueda haber sustitutos psicológicos para las cadenas físicas y los castigos externos. Las tradiciones militares del honor y la lealtad a los camaradas pueden actuar como cadenas invisibles. El estigma de la cobardía, con la desgracia y vergüenza que se le asocia, puede ser tan eficaz como un castigo externo. Además, habida cuenta de la situación hipotética, es racional preferir verse sujeto por estos grilletes psicológicos siempre que nuestros camaradas también lo estén. Si uno tiene que luchar, es mejor, incluso desde un punto de vista puramente egoísta, pertenecer a una unidad disciplinada y con buena moral que ser miembro de una banda de granujas.

No obstante, también está claro que si Tom y Dan pertenecen a una tradición general que les impulsa a mantener sus acuerdos serán capaces de llegar a un resultado prácticamente idéntico. Serían entonces capaces de llegar al pacto de que cada uno se mantuviese en su puesto, y sería entonces la tradición de mantener los acuerdos lo que atase a cada hombre a su posición. Además, dado que es un pacto, cada uno se sentiría obligado a mantenerlo únicamente mientras lo hiciese el otro, y cada uno sabrá que el otro es del mismo parecer. Por tanto, si cada uno puede ver si el otro se mantiene o no en su puesto, cada uno de ellos tendrá un motivo adicional para hacer lo propio, a saber, la certeza de que al quedarse está amando a quedarse a su camarada. Es cierto que, en la práctica, es probable que una tradición general de mantener acuerdos sea bastante menos eficaz en situaciones de peligro extremo (como es el caso de nuestra hipótesis) que las más especializadas tradiciones militares del honor y la lealtad. Sin embargo, tiene la ventaja de ser más flexible. Puede utilizarse como soporte para la realización y el mantenimiento de todo tipo de tratos útiles. Hume tenía toda la razón al decir que un hombre es tanto más útil, lo mismo para sí que para otros, cuanto mayor sea el grado de probidad y honradez con que se le dote.

El particular ejemplo que hemos utilizado aquí para ilustrar esta forma de juego bipersonal tiene la ventaja de ser al mismo tiempo dramático y realista, pero tiene la desventaja de ser poco apto para repetir la prueba con los mismos juga-

dores. Pensemos pues en otro ejemplo, uno en el cual si, por ejemplo, Tom saliese mal parado de su primera decisión aún pudiera sobrevivir para jugar otra vez con Dan. Asumamos que cada hombre manifiesta únicamente una débil tendencia a mantener acuerdos, y que ninguno de ellos puede ver, en ningún caso, si el otro está manteniendo el trato hasta que él mismo se vea obligado bien a mantenerlo bien a romperlo. Añadamos las razonables asunciones de que si ambos hombres mantienen el acuerdo en una ocasión, será más probable que cada uno de ellos lo mantenga la próxima vez, mientras que si uno o ambos lo rompen en una ocasión, será menos probable que cada uno de ellos se atenga al acuerdo en la siguiente oportunidad. Supongamos también que ambos hombres conocen estas tendencias y que cada vez que Tom y Dan practican este juego saben que tendrán que seguir jugándolo uno con el otro en el futuro. Estas asunciones alteran la forma del juego y hacen patente que si, en alguna ocasión, Dan mantiene el acuerdo, será ventajoso para Tom, egoístamente y con miras al porvenir, que él también lo haga, aunque si en otra ocasión Dan lo rompe, será ventajoso para Tom romperlo igualmente. Por supuesto, la situación sigue siendo simétrica. El interés propio ya no urge sin ambigüedad a que cada hombre rompa el acuerdo en todos los casos y por consiguiente sólo se necesitará una débil tendencia a mantener los acuerdos para guardar el equilibrio. Otras asunciones razonablemente obvias y naturales conducen a una conclusión similar si extendemos el juego en otras direcciones y asumimos que existen más de dos jugadores. Será egoístamente ventajoso que cada hombre mantenga el acuerdo si la mayoría de los participantes lo mantiene e intenta reforzarlo, aunque no lo será si la mayoría lo rompe. Éste es, por supuesto, el tipo de relaciones que Hobbes considerara. Una vez más, en cualquier situación de este tipo sólo se necesita una tendencia a mantener los acuerdos razonablemente débil para que el equilibrio se conserve, ya que es una situación que no tiene que superar ningún nítido y unilateral dato de interés propio. Con todo, aún es preciso garantizar el equilibrio: ni siquiera ahora, por sí solo, ningún cálculo racional de interés propio obligará a que cada hombre mantenga el acuerdo.

Sería irrelevante para nuestros propósitos seguir avanzando en la infinita variedad de tipos de situación que podrían es-



tudiarse mediante la teoría de juegos. Sin embargo, es necesario mencionar al menos una complicación. En los ejemplos que hemos puesto hasta ahora, tanto las situaciones iniciales como los acuerdos considerados han sido simétricos, pero es obvio que no tienen por qué serlo necesariamente. A pesar de que Tom y Dan se encuentren al principio en la misma posición, podrían darse varios acuerdos entre ellos, siendo cada uno de esos acuerdos mejor para cada hombre por separado que el resultado, bien de no poder alcanzar acuerdo alguno, bien de romperlo una vez establecido, y eso a pesar de que algunos de dichos acuerdos sean, en diverso nivel, más ventajosos para Tom que para Dan, y viceversa. En esas circunstancias, lo más probable es que el hombre que sea, o tenga apariencia de ser, más reactivo a realizar o a adherirse a un acuerdo, sea también el que obtenga los términos más ventajosos. Aunque la completa intransigencia en las dos partes es desastrosa para ambas, una intransigencia incompleta y relativa es diferencialmente ventajosa para quien la manifieste. Esto es así, como ya he dicho, a pesar de que la situación inicial sea simétrica. Pero si una de las partes tiene menos que perder en caso de no concretar el acuerdo, o menos que ganar si alcanza un acuerdo estable, surgirán nuevas posibilidades de acuerdos desiguales. Los tratos racionales pueden dar como resultado una explotación. Por muy benéficas que sean las cadenas invisibles de las que hemos hablado, nunca pueden constituir una bendición sin mácula.

No hay duda alguna de que las situaciones de la vida real contienen, al menos como parte de su estructura causal relevante, pautas de relación que pueden describirse de manera fecunda mediante varios «juegos» simples y distintos. Una carrera armamentística internacional es un ejemplo obvio. Otro es la situación en que la inflación sólo puede disminuirse si las distintas agrupaciones sindicales pueden llegar al acuerdo de poner límites a sus demandas de incrementos salariales. Uno de los méritos de estos análisis simplificados es que muestran de manera clara el modo en que el resultado conjunto de varias acciones intencionales, incluso las de agentes racionales bien informados, puede ser algo que ninguno de los agentes implicados haya intentado o quisiera intentar lograr. Incluso las ciencias sociales puramente descriptivas se ocupan en buena parte, tal como Popper ha señalado con nitidez, de los efec-

tos indeseados de las acciones intencionales. Sin embargo, desde nuestro punto de vista, el enfoque de la teoría de juegos simplemente refuerza las lecciones que ya habíamos extraído de los argumentos de Protágoras, Hobbes, Hume y Warnock. La moraleja más importante es el valor práctico de la noción de obligación, el valor de un lazo o ligadura invisible y en realidad ficticio. Y esto se aplica tanto si el lazo se presenta bajo la forma de un requisito general que impulse a mantener cualesquiera acuerdos que uno haga como si adopta el aspecto de varias obligaciones específicas como sucede en el caso del honor militar o de la lealtad a los camaradas o a una organización.

Una de las conclusiones que podríamos sentir tentación de extraer de los análisis de la teoría de juegos es que la prudencia no basta, que el cálculo racional del interés propio a largo plazo no es en sí mismo suficiente para hacer que los hombres realicen acuerdos mutuamente benéficos o para obligar a que, una vez alcanzados, los mantengan. Es preciso no obstante tomar algunas precauciones. Esto es cierto en el caso de los «juegos» concretos que hemos descrito. Pero no hay nada en los análisis de la teoría de juegos que elimine la posibilidad de lo que Hobbes sugiere: un acuerdo respaldado por un mecanismo coercitivo (el soberano), una situación en la que los motivos del propio soberano, los de aquellos que le obedecen y los de quienes proporcionan el apoyo que hace posible el poder del soberano mismo, sean todos de tipo exclusivamente prudencial. Es de hecho difícil de ver cómo un constructo semejante puede ser puesto en práctica mediante el operar de motivos egoístas, por muy racional que sea su pauta. Es en cambio menos difícil de comprender cómo, una vez que existe, la fórmula puede mantenerse recurriendo únicamente a este tipo de motivos. La verdadera debilidad de la solución hobbesiana no reside en nada que puedan mostrar los modelos de la teoría de juegos, sino en lo que, precisamente porque son modelos, dejan sin analizar. Las situaciones reales siempre incorporan, además de la estructura desnuda de algún juego razonablemente simple, otras fuerzas o tendencias cuya fuerza varía a lo largo del tiempo. Como he dicho antes, la solución hobbesiana es como un castillo de naipes—cada parte es sostenida por las demás—y, del mismo modo, es también inflexible. Una estructura tendrá mayor tendencia a combarse en respuesta a

las fuerzas cambiantes que la tensan sin por ello derrumbarse si se mantiene unida por lazos entre los que puedan contarse algunos no tan condicionales como los de la prudencia.

##### 5. El contenido del instrumento: ¿conservadurismo o reforma?

En el apartado 3 he hecho algunas consideraciones muy generales sobre la forma del mecanismo al que estamos ahora equiparando la moral en sentido estricto, pero he dicho muy poco de su contenido. De entre los pensadores a los que me he referido, Protágoras, Hume y Warnock se preocupan todos de explicar la cuestión de una moral ya existente, aunque seleccionan y enfatizan partes concretas de la misma. Hobbes escribe como si su tarea fuera la de crear una moral partiendo de cero, aunque la mayoría de las disposiciones que incorpora son de hecho porciones de un cuerpo convencionalmente aceptado de ideas morales. Su condena del castigo retributivo es sólo una excepción parcial, dado que la moral tradicional cristiana contiene sugerencias contradictorias al respecto.

Pero, ¿de verdad es tan obvio que lo que convencionalmente se acepta como moral es justo lo que se precisa? Tal como todos nuestros autores han destacado, el mecanismo de la moral es beneficioso debido a ciertos rasgos contingentes de la condición humana. Pero si son contingentes, también podría suceder que hubiesen cambiado. La contraposición entre Protágoras y Hobbes señala al menos una variación en la escala del problema. Protágoras estaba buscando los principios ordenadores de la ciudad, de la *polis*, y en Grecia una *polis* podía ser francamente pequeña: su problema era averiguar cómo podrían los hombres constituir unidades sociales suficientemente grandes para competir con las bestias salvajes. Pero en el caso de Hobbes el problema era indagar la forma de conseguir un Estado-nación estable. En nuestros días la escala ha vuelto a cambiar: ya no podemos seguir compartiendo la asunción de Hobbes de que únicamente las guerras civiles constituyen una amenaza real y que las guerras internacionales causan relativamente pocos daños. Warnock piensa que es un tanto

impropio que un filósofo quiera dar cualquier explicación de los hechos empíricos contingentes relacionados con las dificultades humanas. Sin embargo, podríamos argumentar que, dado el enfoque general, debería haber dado más explicaciones sobre ellos y no menos.

Los cambios en la situación de los seres humanos que pueden ser relevantes para la moral han sucedido en los últimos ciento cincuenta años, y muy especialmente en los últimos cincuenta años. Aunque son evidentes y bien conocidos, debemos al menos resumirlos. Uno es el incremento de la dependencia internacional mutua. Éste es en parte un hecho económico: una proporción creciente de lo que la gente considera como necesidades se hace posible, directa o indirectamente, gracias a mercancías provenientes de países lejanos. También es parcialmente una cuestión de posibilidades de ayuda: en 1700, los habitantes de Europa no habrían tenido noticias, al menos hasta mucho después, de un terremoto o de una hambruna en la India, e incluso en el caso de que lo hubiesen sabido no hubieran podido hacer nada. Esto ya no es así. Es también, en parte, consecuencia del hecho de que existan movimientos políticos de alcance mundial, así como del hecho de que las guerras locales y los cambios de gobierno puedan tener repercusiones en puntos muy distantes. Una vez más, los avances tecnológicos de diversas clases han puesto en manos de algunos hombres y organizaciones un poder grande y creciente. Esto incluye el poder de hacer daño, por ejemplo, el poder de desencadenar una guerra nuclear. Incluye también el poder para hacer el bien, al menos en apariencia. Tal como dice Belloc:

Antaño, cuando las gentes enfermaban y sufrían rudas pruebas recibían medicamentos de los doctores, y morían; hoy vivimos en una era más feliz, ya que sabemos dos cosas: cómo hacer enfermar a los hombres y cómo mantenerlos así,

de nuevo, el poder puede realizar tanto lo bueno como lo malo, especialmente el poder de producir bienes económicos al coste de una permanente destrucción de los recursos y de provocar cambios radicales en el entorno biológico y físico. Como siempre, los desarrollos en los medios de comunicación (para la información, el entretenimiento, la persuasión o diversas combi-

naciones de los tres) han otorgado a ciertos hombres un poder creciente sobre las mentes de los demás, y probablemente existirá, más pronto que tarde, la posibilidad de aplicar la ingeniería genética a la propia especie humana.

Es tentador referirse a todas estas crecientes capacidades como poderes que la humanidad (o el «hombre») posee en la actualidad y puede utilizar de uno u otro modo o bien abstenerse de hacerlo. Pero esto es completamente desencaminado. La humanidad no es un agente, no tiene unidad de decisión, y no se ve por tanto confrontada a opciones de ninguna clase. Nuestros ejemplos sacados de la teoría de juegos han mostrado aún más claramente lo que ya debería haber estado suficientemente claro sin necesidad de recurrir a ellos: que una pluralidad de agentes racionales que interactúan no constituye en general ningún agente racional, y que el resultado de un cierto número de elecciones no es en general una elección. Los poderes que hemos mencionado están dispersos: constituyen la posesión de algunos hombres, grupos de hombres u organizaciones y son ellos quienes pueden ejercerlos, no «el hombre».

Sin embargo, hay un tipo de cambio que los incrementos del poder tecnológico no han producido. No han logrado falsar la descripción de Hume de la «escasa provisión que la naturaleza ha dispuesto para los deseos [de los hombres]». Ha aumentado lo que los hombres consideran como sus necesidades, en los países económicamente desarrollados, de una forma a nosotros tan rápida como su capacidad para satisfacerlas; y en los países menos desarrollados el crecimiento de la producción ha sido contrarrestado por el aumento de la población.

Puede argumentarse que cambios como éstos, aunque de gran importancia para la filosofía política, tienen poca influencia sobre la moral. Los principios morales descansan sobre la estructura básica y general de los apuros humanos, y esto no cambia. Contra este argumento, es claro que los cambios técnicos han extendido al menos la gama y el alcance de las cuestiones morales. Para nosotros no es tan fácil como lo fue para Hume decir sin pestañear «un indio, es decir, una persona totalmente desconocida para mí». Si estuviéramos considerando la idea de amar a todos nuestros prójimos, ahora serían muchos más los que deberíamos tener en cuenta. Tampoco es tan fácil prescribir una preocupación teórica universal cuando re-

sulta que tenemos cierta capacidad real de llevarla a cabo. Los nuevos poderes hacen surgir embarazosas cuestiones morales en relación con su puesta en práctica. Si podemos mantener a las gentes con vida, o vivas a medias, con determinados costes, la cuestión de si debemos o no hacerlo no es ociosa. Las cuestiones que se plantean en torno a las respectivas demandas de las generaciones presentes y futuras ya no son puramente académicas, puesto que las cosas que hacemos ahora pueden afectar radicalmente, de modos no totalmente impredecibles, a las expectativas de las generaciones futuras. El control del número de habitantes, de ciertos países al menos, y, como posibilidad más remota, de todo el mundo, se vuelve poco a poco un objetivo político concebible. Estos hechos, unidos a los riesgos de la manipulación psicológica y de la ingeniería genética, introducen una desconcertante circularidad en la relación entre los principios morales y el bienestar humano. La especie humana ha dejado de ser algo determinado cuyos miembros tienen intereses razonablemente fijos y cuyo grado de satisfacción es la pauta que permite medir el nivel de bienestar y sopesar las decisiones morales. Algunas de las decisiones que se plantean a la valoración moral pueden determinar por sí mismas cuáles hayan de ser esos intereses e incluso en qué haya de convertirse la propia especie humana.

Al comienzo de este capítulo he dicho que la moral no puede ser descubierta sino que debe realizarse. No podemos deshacernos de esto de un plumazo añadiendo «pero es algo ya realizado hace mucho tiempo». Bien podría suceder que necesitara ser rehecha en parte. Por supuesto, sólo en parte. Nada ha alterado ni alterará la importancia de ser capaz de hacer y mantener acuerdos y confiar asimismo en que otros los mantengan. La tercera ley de la naturaleza de Hobbes, que los hombres realicen los acuerdos que establezcan, es un fragmento eterno e inmutable de la moral. Sin embargo, algunas obligaciones de carácter más específico y tradicionalmente vinculadas a la posición, es decir, no generadas por contrato, son prescindibles. La utilidad del patriotismo, por ejemplo, puede haber caducado. Si no podemos estar completamente satisfechos con el contenido de lo que podemos reconocer como corriente principal de la tradición moral, deberemos sentirnos menos satisfechos todavía con la alternativa de dejar que el entero contenido de la moral sea determinado por la conciencia o sentido moral de

cada agente individual. Como dice Anscombe, «La conciencia de un hombre puede sugerirle que haga las cosas más viles», y esto no puede sorprendernos cuando caemos en la cuenta de cómo se forman y cuál es el operar de nuestras conciencias. Su actuación se basa en la incorporación durante la infancia, en el seno de las mentes individuales, de exigencias morales externas modificadas no obstante, a menudo inconscientemente, por muchos motivos, y, por consiguiente, las conciencias pueden servir de cauce para deseos que de otro modo quedarían reprimidos. Incluso en el caso de que nos pusieramos en la perspectiva más optimista posible y asumiésemos que, en general, las conciencias de los hombres han sido moldeadas adecuadamente por las fuerzas evolutivas, lo más que podríamos esperar es que hubieran dejado sentados algunos principios que se *hubiesen revelado* útiles. A diferencia del Dios al que ha reemplazado, no puede suponerse que la selección natural posea o incorpore ninguna presciencia.

Esto no implica negar que el sentido y las virtudes morales sean partes esenciales de la *forma* del mecanismo de la moral. Dado que la prudencia no es suficiente, incluso en el caso de que la combinemos, en la solución hobbesiana, con un mecanismo coercitivo, es importante que haya una amplia tendencia a actuar sobre bases morales. Lo que sí estoy negando en cambio, y es algo que lo anterior no implica, es que el mejor *contenido* posible para el mecanismo de la moral nos lo proporcione el sentido moral específico que cada agente haya podido adquirir.

## VI

### El utilitarismo

#### 1. El utilitarismo de los actos

Al comienzo del capítulo 5, la *moral en sentido estricto* era rudimentariamente caracterizada como un particular tipo de restricciones sobre la conducta. En el resto de ese capítulo se estudiaba por qué eran necesarias tales restricciones para el florecimiento de la vida humana. La perspectiva de la moral en sentido estricto que allí esbozábamos era por consiguiente utilitarista en el muy amplio sentido de considerar el bienestar humano en general como, de algún modo, el fundamento de la moral. Con todo, usar el término «utilitarismo» de forma tan amplia no es demasiado esclarecedor. Es mejor restringirlo, o restringir algunas cualificadas variantes del mismo, hasta lograr perspectivas más específicas acerca del modo en que las conclusiones morales pueden derivarse de, o fundarse sobre, la felicidad humana. En definitiva es mejor restringirlo hasta obtener métodos específicos para determinar el contenido del sistema moral de primer orden.

Una de estas perspectivas es el utilitarismo extremo o utilitarismo de los actos. Este punto de vista mantiene que allí donde un agente tenga una elección entre dos o más cursos de acción (o inacción) el acto correcto es aquel que produzca la mayor felicidad, no simplemente para el propio agente sino