

activas? No obstante, el papel de la distinción entre activo y pasivo en la vida humana es omnipresente y profundo. La diferencia entre pasividad y actividad es inescindible del hecho de que existimos como seres y agentes, y no sólo como escenarios en los que, por casualidad, ocurren ciertos acontecimientos.

Aunque no me he ocupado de analizar esta distinción como tal, figura de diversas maneras en los numerosos análisis de la necesidad y la voluntad que transmiten otra preocupación temática de mis ensayos. Me parece que descubrir qué somos es, fundamentalmente —aunque sólo quizá entre otras cuestiones—, descubrir qué debemos ser. Y, por ello, en la medida en que una persona se identifica por su voluntad, es cuestión de descubrir lo que no podemos evitar desear o lo que no podemos lograr desear. La noción de que la necesidad no debilita en forma inevitable la autonomía es conocida y ampliamente aceptada. Sin embargo, la necesidad no sólo es compatible con la autonomía; en ciertos aspectos, le es esencial. Debe haber límites a nuestra libertad si hemos de tener una realidad personal suficiente para ejercer la auténtica autonomía. Lo que no tiene límites, no tiene forma. Del mismo modo, una persona no puede tener una identidad o una naturaleza esencial como agente a menos que esté obligada con respecto a esa característica de sí misma —es decir, la voluntad—, cuya forma coincide, fielmente, con lo que es esa persona, y que también lo da a conocer. Varios de mis ensayos —entre ellos el último, que hasta ahora no había sido publicado— se refieren a los intentos de arrojar luz sobre la creación de estas necesidades y sobre la importancia que tienen para nosotros.

## 1 Posibilidades alternativas y responsabilidad moral

Un principio al que llamaré el *principio de las posibilidades alternativas* ha desempeñado un papel dominante en casi todas las investigaciones recientes acerca del problema del libre albedrío. Este principio afirma que una persona es moralmente responsable de lo que ha hecho sólo en el caso de que hubiera podido comportarse de otra manera. Su significado exacto es un tema de controversia, en particular con respecto a la cuestión de si alguien que lo acepta queda comprometido a creer que la responsabilidad moral y el determinismo son incompatibles. Sin embargo, prácticamente nadie parece inclinado a negar o siquiera a cuestionar que el principio de las posibilidades alternativas (interpretado de una u otra manera) sea verdad. En general, siempre pareció tan abrumadoramente razonable que algunos filósofos han llegado a caracterizarlo como una verdad *a priori*. Es evidente que las personas cuyas concepciones de libre albedrío o de responsabilidad moral están radicalmente enfrentadas encuentran en él un terreno común firme y conveniente sobre el que pueden adoptar con provecho sus posturas opuestas.

Pero el principio de las posibilidades alternativas es falso. Una persona bien puede ser moralmente responsable de lo que hizo aunque no haya podido comportarse de otra manera. La verosimilitud del principio es una ilusión, que puede desvanecerse si centramos la atención en los fenómenos morales pertinentes.

## I

Cuando se busca ilustrar el principio de las posibilidades alternativas, es muy natural pensar en situaciones en las que las mismas circunstancias provocan que una persona haga algo y, a la vez, hacen imposible que evite hacerlo. Éstas incluyen, por ejemplo, situaciones en las que una persona es coaccionada a hacer algo o es impelida de actuar por una sugestión hipnótica, o en las que un impulso interno la lleva a actuar como lo hace. En situaciones de esta índole, hay circunstancias que impiden que la persona actúe de otra manera, y estas mismas circunstancias también sirven para provocar que haga lo que hace.

No obstante, puede haber circunstancias que constituyen condiciones suficientes para que alguien lleve a cabo una acción y que, por lo tanto, impiden que la persona actúe de otra forma, pero que, en realidad, no obligan a la persona a actuar ni a producir su acción de ninguna manera. Una persona puede hacer algo en circunstancias que no le dejan más alternativa que hacerlo, sin que estas circunstancias en realidad la muevan o la conduzcan a hacerlo, sin que desempeñen ningún papel, en efecto, en provocar que actúe como lo hace.

Un examen de situaciones caracterizadas por circunstancias de esta clase arroja dudas, creo yo, sobre la relevancia del hecho de que una persona que ha actuado de una manera no podría haber actuado de otra, respecto de cuestiones de responsabilidad moral. Propongo desarrollar algunos ejemplos de esta clase en el contexto de un debate sobre la coacción y sugerir que nuestras intuiciones morales con respecto a estos ejemplos tienden a desconfirmar el principio de las posibilidades alternativas. Luego discutiré el principio en términos más generales, explicaré lo que pienso que son sus errores y describiré brevemente y sin argumento cómo podría revisarse en forma apropiada.

## II

En general, estamos de acuerdo en que una persona que ha sido coaccionada a hacer algo no lo hizo libremente y no es moralmente responsable de haberlo hecho. Ahora bien, podría parecer que la doctrina de que la coacción y la responsabilidad moral son mutuamente exclusivas no es más que una versión algo particularizada del principio de las posibilidades alternativas. Es natural decir que alguien que ha sido coaccionado a hacer algo no podía hacer otra cosa. Y fácilmente podría parecer que el hecho de ser coaccionado priva a alguien de la libertad y la responsabilidad moral, simplemente porque es un caso especial de imposibilidad de hacer otra cosa. De esta manera, el principio de las posibilidades alternativas puede obtener cierta credibilidad de su asociación con la muy razonable propuesta de que la responsabilidad moral queda excluida por la coacción.

Sin embargo, no está bien que así sea. El hecho de que una persona haya sido coaccionada a actuar como lo hizo puede implicar tanto que no podía hacer otra cosa como que no es responsablemente moral de su acción. Pero su falta de responsabilidad moral no es consecuencia del hecho de no haber podido hacer otra cosa. En otras palabras, no es correcto interpretar la doctrina de que la coacción excluye la responsabilidad moral como una versión particularizada del principio de las posibilidades alternativas.

Supongamos que alguien es amenazado en forma convincente con un castigo que considera inaceptable y que luego hace lo que la persona que lo amenazó le pide. Podemos imaginarnos detalles que nos llevarían a pensar que la persona fue coaccionada para llevar a cabo la acción en cuestión, que no pudo hacer otra cosa y que no es moralmente responsable de haber hecho lo que hizo. No obstante, ¿qué tienen las situaciones de esta clase que garantizan el juicio de que la persona amenazada no es moralmente responsable de su acto?

Esta pregunta puede enfocarse considerando situaciones como la siguiente: Jones tiene razones para decidir hacer algo, luego alguien lo amenaza con un castigo muy severo (tan severo que cualquier persona razonable se rendiría a la amenaza) a menos que haga precisamente aquello, y Jones lo hace. ¿Responsabilizaremos moralmente a Jones por lo que hizo? Creo que eso dependerá de los pa-

peles que creemos que fueron desempeñados, al llevarlo a actuar, por su decisión original y por la amenaza.

Una posibilidad es que Jones<sub>1</sub> no sea un hombre razonable; más bien, es un hombre que hace lo que decidió hacer sin importar lo que suceda después y sin importar el costo. En ese caso, la amenaza en realidad no ejerció una fuerza efectiva sobre él. Actuó sin pensar en ella, casi como si no fuera consciente de su existencia. Si, en efecto, así fue, la situación no implicó ninguna coacción. La amenaza no llevó a Jones<sub>1</sub> a hacer lo que hizo. Como tampoco fue, de hecho, suficiente para impedir que hiciera otra cosa: si su decisión anterior hubiera sido la de hacer otra cosa, la amenaza no lo habría disuadido en absoluto. Parece evidente que, en estas circunstancias, el hecho de que Jones<sub>1</sub> fuera amenazado no reduce de ningún modo la responsabilidad moral que tendría, de otra manera, por su acto. Este ejemplo, sin embargo, no es un contraejemplo de la doctrina que afirma que la coacción libera de responsabilidad ni del principio de las posibilidades alternativas. Pues supusimos que Jones<sub>1</sub> es un hombre sobre quien la amenaza no tiene ningún efecto coercitivo y, por ello, que en realidad no lo privó de alternativas a lo que hizo.

Otra posibilidad es que Jones<sub>2</sub> fuera empujado a tomar una decisión precipitada por la amenaza. Dada esa amenaza, habría llevado a cabo la acción sin importar qué decisión había tomado. Además, la amenaza lo perturbó tanto que se olvidó por completo de su decisión anterior e hizo todo lo que se le pedía porque tenía terror del castigo con el que había sido amenazado. En este caso, el hecho de que ya hubiera decidido llevarla a cabo no pesa sobre su realización de la acción. Cuando cayó en la cuenta, no pensó en otra cosa que en la amenaza, y sólo el temor lo llevó a actuar. El hecho de que anteriormente Jones<sub>2</sub> hubiera decidido, por sus propios motivos, actuar de esa manera puede ser relevante para una evaluación de su carácter; puede tener toda la responsabilidad moral para haber tomado *esa* decisión. Pero no se puede decir que sea moralmente responsable de su acto, puesto que llevó a cabo la acción simplemente como resultado de la coacción a la que fue sometido. Su decisión anterior no desempeñó ningún papel en provocar que hiciera lo que hizo y, por lo tanto, sería gratuito asignarle un papel en la evaluación moral de su acción.

Ahora consideremos una tercera posibilidad. Jones<sub>3</sub> no se sintió empujado por la amenaza a tomar una decisión precipitada y tampoco fue indiferente a ella. La amenaza lo impresionó, como impresionaría a cualquier hombre razonable, y él se habría sometido a ella sin reservas si no hubiera tomado ya una decisión que coincidía con la que se le exigía. Sin embargo, de hecho, él llevó a cabo la acción en cuestión sobre la base de la decisión que había tomado antes de recibir la amenaza. Cuando actuó, en realidad no estaba motivado por la amenaza sino únicamente por las consideraciones que lo habían llevado a actuar, aunque sí lo habría estado de no haber tenido suficientes motivos para llevar a cabo la acción en cuestión.

Sin duda, en un caso como éste será muy difícil para cualquiera saber qué pasó exactamente. ¿Acaso Jones<sub>3</sub> llevó a cabo la acción en virtud de la amenaza, o sus razones para actuar fueron simplemente las que ya lo habían persuadido de hacerlo? ¿O actuó sobre la base de dos motivos, cada uno de los cuales fue suficiente para su acción? Sin embargo, no es imposible que la situación fuera más clara de lo que son en general las situaciones de esta clase. Y supongamos que para nosotros es obvio que Jones<sub>3</sub> actuó sobre la base de su propia decisión y no por la amenaza. Entonces pienso que estaría justificado considerar que su responsabilidad moral por lo que hizo no está afectada por la amenaza aunque, dado que de todas maneras se habría rendido ante la amenaza, no podía haber evitado lo que hizo. Para nosotros sería completamente razonable emitir el mismo juicio con respecto a su responsabilidad que el que habríamos emitido si no hubiéramos tenido conocimiento de la amenaza, puesto que la amenaza de hecho no ejerció influencia alguna sobre su realización de la acción. Él hizo lo que hizo de la misma manera que si no hubiera recibido la amenaza.

### III

A primera vista, el caso de Jones<sub>3</sub> parecería combinar coacción con responsabilidad moral y, por lo tanto, proporcionar un contraejemplo para la doctrina de que la coacción libera de responsabili-

dad. En realidad, no es tan seguro que lo haga, no obstante, porque no queda claro si el ejemplo constituye una instancia genuina de coacción. ¿Podemos decir que Jones<sub>3</sub> fue coaccionado a hacer algo, cuando ya había decidido hacerlo por su cuenta y cuando lo hizo enteramente sobre la base de esa decisión? ¿O sería más correcto decir que Jones<sub>3</sub> no fue coaccionado a hacer lo que hizo aunque él mismo reconociera que sentía una fuerza irresistible en virtud de la cual debía hacerlo? Mis propias intuiciones lingüísticas me conducen hacia la segunda alternativa, pero son algo equívocas. Quizá podamos decir cualquiera de estas dos cosas, o quizá debamos añadir una explicación calificativa a cualquiera de las que digamos.

Sin embargo, esta oscuridad no nos impide extraer una importante moraleja de un análisis del ejemplo. Supongamos que decidimos decir que Jones<sub>3</sub> *no* fue coaccionado. Nuestra base para decir esto será claramente que es incorrecto considerar que un hombre fue coaccionado a hacer algo a menos que lo haga *a causa de* la fuerza coercitiva ejercida contra él. El hecho de que haya recibido una amenaza irresistible, entonces, no implicará que la persona que la recibe es coaccionada a hacer lo que hace. También será necesario que la amenaza sea la verdadera razón para hacerlo. Por otro lado, supongamos que decidimos decir que Jones<sub>3</sub> *fue* coaccionado. Entonces seguramente admitiremos que ser coaccionado no excluye ser moralmente responsable. Y lo más probable es que lleguemos a la perspectiva de que la coacción afecta la responsabilidad moral de una persona sólo cuando la persona actúa como lo hace porque está coaccionada a hacerlo, por ejemplo, cuando el hecho de que ha sido coaccionada da cuenta de su acción.

Sea lo que fuere que decidamos decir, entonces, reconoceremos que la doctrina de que la coacción excluye la responsabilidad moral no es una versión particularizada del principio de las posibilidades alternativas. Las situaciones en las que una persona que hace algo no puede hacer otra cosa porque está sometida a un poder coercitivo o bien no son en absoluto instancias de coacción o bien son situaciones en las que la persona aún puede ser moralmente responsable de lo que hace si no lo hiciera por la coacción. Cuando disculpamos a una persona que ha sido coaccionada, no la disculpamos porque no pudo hacer otra cosa. Aunque alguien sea some-

tido a una fuerza coercitiva que descarta la posibilidad de que lleve a cabo otra acción, puede, a pesar de todo, tener responsabilidad moral por realizar esa acción.

#### IV

En la medida en que el principio de las posibilidades alternativas deriva su verosimilitud de la asociación con la doctrina de que la coacción excluye la responsabilidad moral, una clara comprensión de esta última disminuye el atractivo de la primera. Sin duda, el caso de Jones<sub>3</sub> puede, al parecer, hacer algo más que iluminar la relación entre las dos doctrinas. Quizá también parecería brindar un contraejemplo decisivo del principio de las posibilidades alternativas y, por tanto, mostrar que el principio es falso, puesto que el carácter irresistible de la amenaza a la que Jones<sub>3</sub> es sometido bien podría interpretarse en el sentido de que él no puede más que realizar la acción que realiza. Y, sin embargo, la amenaza, ya que Jones<sub>3</sub> lleva a cabo la acción sin tenerla en cuenta, no reduce su responsabilidad moral por lo que hace.

Sin duda, se planteará la siguiente objeción contra la sugerencia de que el caso de Jones<sub>3</sub> es un contraejemplo del principio de las posibilidades alternativas. Quizás, en cierto sentido, Jones<sub>3</sub> no puede realizar otra acción más que la que hace, ya que es un hombre razonable, y la amenaza a la que se enfrenta es suficiente para movilizar a cualquier hombre razonable. Pero no es este sentido el que guarda relación con el principio de las posibilidades alternativas. Su conocimiento de que va a sufrir un castigo intolerable y severo no significa que Jones<sub>3</sub>, estrictamente hablando, *no pueda* llevar a cabo ninguna otra acción que no sea la que, en efecto, lleva a cabo. Después de todo, aún le queda abierta, y esto es crucial, la posibilidad de desafiar la amenaza si desea hacerlo y aceptar el castigo que su acción le provocaría. En el sentido en que el principio de las posibilidades alternativas emplea el concepto de “podría haber hecho otra cosa”, la incapacidad de Jones<sub>3</sub> de resistirse a la amenaza no significa que no pueda hacer otra cosa que llevar a cabo la

acción que realiza. De allí que el caso de Jones<sub>3</sub> no constituya un ejemplo contrario al principio.

No propongo considerar en qué sentido el concepto de “podría haber hecho otra cosa” figura en el principio de las posibilidades alternativas, como tampoco intentaré medir la fuerza de la objeción que acabo de describir.<sup>1</sup> Porque considero que la fuerza que se piense que pueda tener esta objeción puede ser desviada alterando el ejemplo de la siguiente manera.<sup>2</sup> Supongamos que alguien —digamos Black— quiere que Jones<sub>4</sub> lleve a cabo cierta acción. Black está dispuesto a llegar lejos para salirse con la suya, pero prefiere no mostrar sus intenciones innecesariamente. De modo que espera hasta que Jones<sub>4</sub> esté a punto de tomar una decisión sobre qué hacer, y no hace nada a menos que le resulte claro (Black es un excelente juez de esas cosas) que Jones<sub>4</sub> va a decidirse a hacer algo que *no* sea lo que él quiere que haga. Si resulta claro que Jones<sub>4</sub> va a decidirse a hacer otra cosa, Black toma medidas efectivas para asegurarse de que Jones<sub>4</sub> se decida a hacer, y que haga, lo que él quiere que haga.<sup>3</sup> Cualesquiera sean las preferencias e inclinaciones iniciales de Jones<sub>4</sub>, entonces, Black se saldrá con la suya.

- 1 Los dos conceptos principales empleados en el principio de posibilidades alternativas son “moralmente responsable” y “podría haber hecho otra cosa”. Discutir el principio sin analizar alguno de estos conceptos parecería un intento de piratería. El lector debería notar que mi bandera con la calavera y los huesos ha sido desplegada.
- 2 Después de pensar en el ejemplo que estoy a punto de desarrollar, me enteré de que Robert Nozick, en unas conferencias que dio hace varios años, había formulado un ejemplo del mismo tipo general y lo había propuesto como un contraejemplo del principio de posibilidades alternativas.
- 3 La suposición de que Black puede predecir lo que Jones<sub>4</sub> va a decidir no da por sentada la cuestión del determinismo. Podemos imaginar que Jones<sub>4</sub> muchas veces se ha enfrentado a las alternativas A y B a las que ahora se enfrenta, y que su gesto invariablemente se frunció cuando estaba a punto de llevar a cabo A y nunca cuando estaba por decidirse a hacer B. Sabiendo esto y observando el gesto, Black tendría una base para la predicción. Seguramente, esto supone que existe una cierta relación causal entre el estado de Jones<sub>4</sub> en el momento del gesto fruncido y sus estados subsiguientes. Sin embargo, cualquier perspectiva razonable de decisión o de acción admitirá que tomar una decisión y llevar a cabo una acción suponen etapas anteriores y posteriores, con relaciones causales entre ellas, y a tal punto que las etapas anteriores no son parte de la decisión ni de la acción. El ejemplo no requiere que estas etapas anteriores estén relacionadas, desde un punto de vista determinista, con sucesos aun anteriores.

¿Qué medidas tomará Black, si cree que debe tomar medidas, a fin de asegurar que Jones<sub>4</sub> decida y actúe como él desea? Cualquiera que tenga una teoría con respecto a lo que significa “podría haber hecho otra cosa” está en condiciones de responder a esta pregunta describiendo las medidas que consideraría suficientes para garantizar que, en el sentido pertinente, Jones<sub>4</sub> no puede hacer otra cosa. Que Black pronuncie una terrible amenaza y, de esta manera, obligue a Jones<sub>4</sub> a llevar a cabo la acción deseada y, a su vez, le impida realizar una prohibida. Que Black le dé a Jones<sub>4</sub> una poción o lo hipnotice, y de alguna otra manera similar, genere en Jones<sub>4</sub> una compulsión interior irresistible de llevar a cabo el acto que Black desea y de evitar otros. O que Black manipule los procesos mínimos del cerebro y el sistema nervioso de Jones<sub>4</sub> de alguna manera más directa, para que las fuerzas causales que entran y salen de sus sinapsis y pasan por los nervios del pobre hombre determinen que elija un acto y que actúe de una manera y no de otra. En otras palabras, cualesquiera sean las condiciones en las que se sostenga que Jones<sub>4</sub> no puede hacer otra cosa, que Black haga prevalecer esas condiciones. La estructura del ejemplo es bastante flexible, creo, para hallar una salida a cualquier cargo de irrelevancia acomodando la doctrina sobre la que se basa el cargo.<sup>4</sup>

Ahora supongamos que Black nunca tiene que intervenir porque Jones<sub>4</sub>, por razones propias, decide llevar a cabo, y lo hace, la acción misma que Black quiere que lleve a cabo. En ese caso, parece claro, Jones<sub>4</sub> tendrá precisamente la misma responsabilidad moral por lo que hace que la que habría tenido si Black no hubiera estado dispuesto a tomar medidas para asegurar que lo hiciera. Sería muy poco razonable excusar a Jones<sub>4</sub> por su acción, o negar el elogio al que ésta normalmente le daría derecho, sobre la base de que no podía haber hecho otra cosa. Esta situación no lo llevó en absoluto a

- 4 El ejemplo también es lo suficientemente flexible para permitir la eliminación por completo de Black. Cualquiera que crea que la eficacia del ejemplo se debilita por su dependencia de un manipulador humano que impone su voluntad a Jones<sub>4</sub> puede sustituir a Black por una máquina programada para hacer lo que hace Black. Si esto aun no basta, olvidémonos de Black y de la máquina y supongamos que su papel es desempeñado por fuerzas naturales que no involucran ninguna voluntad ni plan en absoluto.

actuar como lo hizo. Habría actuado de la misma manera aunque ello no hubiera ocurrido. En efecto, todo sucedió como habría sucedido sin la presencia de Black en esa situación y sin su disposición para intervenir en ella.

En este ejemplo, hay suficientes condiciones para que Jones<sub>4</sub> lleve a cabo la acción en cuestión. La acción que realice no depende de él. Por supuesto, en cierta manera depende de él si actúa por su propia cuenta o como resultado de la intervención de Black. Ello depende de la acción que él mismo esté inclinado a realizar. Pero ya sea que finalmente actúe por su cuenta o como resultado de la intervención de Black, lleva a cabo la misma acción. No tiene más alternativa que hacer lo que Black quiere que haga. Sin embargo, si lo hace por su cuenta, su responsabilidad moral por hacerlo no se ve afectada por el hecho de que Black estuviera acechando con intenciones siniestras, ya que estas intenciones nunca entran en juego.

v

El hecho de que una persona no hubiera podido evitar hacer algo es una condición suficiente para que lo haya hecho. Sin embargo, como muestran algunos de mis ejemplos, es posible que este hecho no desempeñe ningún papel en la explicación de por qué lo hizo. Es posible que no figure en absoluto entre las circunstancias que en realidad provocaron que hiciera lo que hizo, de modo que su acción debe ser explicada sobre otra base por completo diferente. Es decir, aunque la persona hubiera sido incapaz de hacer otra cosa, podría no ser el caso que actuara como lo hizo *porque* no pudo hacer otra cosa. Ahora bien, si alguien no tuvo alternativas para llevar a cabo cierta acción, pero no la realizó porque no pudo hacer otra cosa, entonces habría llevado a cabo exactamente la misma acción aunque *hubiera podido* hacer otra cosa. Las circunstancias que le impidieron hacer otra cosa se podrían haber extraído de la situación sin afectar lo que pasó o por qué pasó en cualquier caso. La razón que efectivamente ha conducido a la persona a hacer lo que hizo o que la obligó

a hacerlo la habría conducido u obligado a hacerlo aun en el caso en que esa persona hubiera podido hacer otra cosa.

Así, no habría habido diferencia, en lo que concierne a su acción o a cómo llegó a realizarla, si las circunstancias que le hicieron imposible evitar realizarla no hubieran prevalecido. Es claro que el hecho de que no haya podido hacer otra cosa no provee ninguna base para suponer que *podría* haber hecho otra cosa de haber sido capaz. Cuando un hecho es, de esta manera, irrelevante respecto del problema de dar cuenta de la acción de una persona, parece infundado asignarle algún peso en la evaluación de su responsabilidad moral. ¿Por qué se debería considerar el hecho al emitir un juicio moral con respecto a la persona cuando no ayuda en absoluto a comprender qué la llevó a actuar como lo hizo o qué podría haber hecho en otras circunstancias?

Ésta, entonces, es la razón por la que el principio de las posibilidades alternativas es erróneo. Afirma que una persona no tiene responsabilidad moral —es decir, que está disculpada— por haber llevado a cabo una acción si hubo circunstancias que le hicieron imposible evitar realizarla. Sin embargo, quizás haya circunstancias que hagan imposible que una persona evite realizar una acción sin que esas circunstancias provoquen en modo alguno que realice esa acción. Ciertamente no sería bueno que la persona se refiriera a circunstancias de esta clase en un intento por absolverse de la responsabilidad moral por llevar a cabo la acción en cuestión. Ello se debe a que, de acuerdo con nuestra hipótesis, en realidad esas circunstancias no tuvieron nada que ver con que la persona haya hecho lo que hizo. Habría hecho exactamente lo mismo y habría sido conducida u obligada a hacerlo de la misma manera, aunque aquéllas no hubieran prevalecido.

Por cierto, con frecuencia disculpamos a las personas por lo que han hecho cuando nos dicen (y les creemos) que no podían hacer otra cosa. Pero esto es porque suponemos que lo que nos dicen sirve para explicar por qué hicieron lo que hicieron. Damos por sentado que no están faltando a la sinceridad, como estaría haciendo alguien que adujera como excusa el hecho de que no tenía modo de evitar hacer lo que hizo, pero que supiera perfectamente que ésta no fue, de ningún modo, la razón por la que lo hizo.

Lo que dije puede sugerir que el principio de las posibilidades alternativas debería revisarse como para afirmar que una persona no es moralmente responsable de lo que ha hecho si lo hizo porque no podía hacer otra cosa. Se observará que esta revisión del principio no afecta seriamente los argumentos de quienes confiaron en el principio original en un intento por sostener que la responsabilidad moral y el determinismo son incompatibles, puesto que si existió determinación causal para que una persona realizara una acción, será verdad que la persona la realizó en virtud de esos determinantes causales. Y si el hecho de que la acción de una persona estuviera determinada causalmente significa que la persona no podía hacer otra cosa, como en general suponen los filósofos que defienden la tesis de la incompatibilidad, entonces el hecho de que la acción de una persona estuviera determinada causalmente significará que la persona la realizó porque no podía hacer otra cosa. El principio revisado de las posibilidades alternativas implicará, en cuanto a esta suposición con respecto al significado de “podía haber hecho otra cosa”, que una persona no es moralmente responsable de lo que hizo si su acción estuvo determinada causalmente. No obstante, no creo que esta revisión del principio sea aceptable.

Supongamos que una persona nos dice que hizo lo que hizo porque no podía hacer otra cosa; o supongamos que hace una afirmación similar: que hizo lo que hizo porque tenía que hacerlo. Con frecuencia aceptamos enunciados como éstos (si los creemos) como excusas válidas, y esos enunciados, a primera vista, parecen invocar el principio revisado de las posibilidades alternativas. Pero creo que cuando aceptamos enunciados de ese tipo como excusas válidas es porque suponemos que nos dicen más de lo que los enunciados transmiten estricta y literalmente. Entendemos que la persona que ofrece la excusa quiere decir que hizo lo que hizo *sólo porque* no podía hacer otra cosa o *sólo porque* tenía que hacerlo. Y entendemos que quiere decir, más en particular, que cuando hizo lo que hizo no fue porque eso fuera lo que realmente quería hacer. El principio de las posibilidades alternativas, entonces, debería ser reemplazado, en mi opinión, por el siguiente principio: una persona no es moralmente responsable de lo que hizo si lo hizo sólo

porque no podía hacer otra cosa. Este principio no parece estar en conflicto con la visión de que la responsabilidad moral es compatible con el determinismo.

Lo siguiente puede ser verdad: existieron circunstancias que hicieron imposible que una persona evitara hacer algo; estas circunstancias, en realidad, desempeñaron un papel para provocar que esa persona hiciera lo que hizo, de modo que es correcto decir que lo hizo porque no podía hacer otra cosa; la persona realmente quería hacer lo que hizo; lo hizo porque era lo que realmente quería hacer, de modo que no es correcto decir que hizo lo que hizo sólo porque no podía hacer otra cosa. En estas condiciones, la persona bien puede ser moralmente responsable de lo que hizo. Por otro lado, no será moralmente responsable de lo que hizo si lo hizo sólo porque no podía hacer otra cosa, aunque lo que hizo fuera algo que realmente quería hacer.

## 2

### La libertad de la voluntad y el concepto de persona

Lo que los filósofos han aceptado últimamente como análisis del concepto de persona no es, en absoluto, el análisis de *ese* concepto. Strawson, cuya manera de emplearlo representa la pauta generalizada en la actualidad, identifica el concepto de persona con “el concepto de un tipo de entidad tal que *tanto* los predicados que le atribuyen estados de conciencia *como* los predicados que le atribuyen características corpóreas [...] pueden ser igualmente aplicados a un individuo único de ese tipo único”.<sup>1</sup> Sin embargo, hay muchas entidades aparte de las personas que tienen propiedades tanto mentales como físicas. Da la casualidad —aunque parece extraordinario que así sea— de que no hay una palabra común en inglés para el tipo de entidad que Strawson tiene en mente, un tipo que incluye no sólo a seres humanos, sino también a animales de diversas especies menores. No obstante, ello no justifica el mal uso de un término filosófico valioso.

La cuestión de si los miembros de algunas especies animales son personas no va a ser resuelta, seguramente, mediante la mera determinación de si es correcto emplear para ellos, además de

<sup>1</sup> P. F. Strawson, *Individuals*, Londres, Methuen, 1959, pp. 101-102 [trad. esp.: *Individuos*, Madrid, Taurus, 1989.] El uso que Ayer hace de “persona” es similar: “Es característico de las personas, en este sentido, que aparte de tener diversas propiedades físicas [...] también se les reconozcan diversas formas de conciencia”, en A. J. Ayer, *The Concept of a Person*, Nueva York, St. Martin's, 1963, p. 82 [trad. esp.: *El concepto de persona*, Barcelona, Seix Barral, 1966]. Lo que les preocupa a Strawson y a Ayer es el problema de comprender la relación entre la mente y el cuerpo, y no el problema, bastante diferente, de comprender qué significa ser una criatura que no sólo tiene una mente y un cuerpo, sino que también es una persona.

predicados que les atribuyan características corpóreas, predicados que les atribuyan estados de conciencia. De hecho, atenta contra nuestro lenguaje emplear el término “persona” para aquellas numerosas criaturas que tienen tanto propiedades psicológicas como materiales, pero que, evidentemente, no son personas en ningún sentido normal de la palabra. Este mal uso del lenguaje es, sin duda, ajeno a cualquier error teórico. Sin embargo, aunque la ofensa es “meramente verbal”, causa un daño importante, debido a que reduce en forma gratuita nuestro vocabulario filosófico y aumenta la probabilidad de que pasemos por alto la importante área de investigación con la que más naturalmente se asocia el término “persona”. Se podría haber esperado que para los filósofos ningún problema presentara un mayor y más constante interés que el de comprender lo que somos en esencia. No obstante, por lo general se descuida tanto este problema que es habitual llevarse el nombre por delante casi sin que se note y, evidentemente, sin provocar ningún sentimiento generalizado de pérdida.

Hay un sentido en que la palabra “persona” es simplemente la forma singular de “gente” y en que ambos términos sólo connotan la pertenencia a cierta especie biológica. Sin embargo, en los sentidos de la palabra que son de mayor interés filosófico, los criterios para ser una persona no sirven fundamentalmente para distinguir a los miembros de nuestra propia especie de los miembros de otras especies. Por el contrario, están formulados para captar aquellos atributos que son el centro de nuestra preocupación más humana con nosotros mismos y la fuente de lo que consideramos más importante y más problemático en nuestras vidas. Ahora bien, estos atributos tendrían igual importancia para nosotros incluso si no fueran, de hecho, peculiares y comunes a los miembros de nuestra propia especie. Lo que más nos interesa de la condición humana no nos interesaría menos si también fuera un rasgo característico de la condición de otras criaturas.

Nuestro concepto de nosotros como personas no se puede comprender, por tanto, como un concepto de atributos que necesariamente son específicos de la especie. Es posible, desde el punto de vista conceptual, que miembros de especies no humanas nuevas —o

incluso conocidas— sean personas; y también es posible, desde el mismo punto de vista, que algunos miembros de la especie humana no sean personas. En realidad suponemos, por otra parte, que ningún miembro de otra especie es una persona. De acuerdo con esto, se presume que las personas se definen esencialmente por una serie de características que —ya sea correcta o incorrectamente— en general suponemos específicamente humanas.

Creo que una diferencia esencial entre las personas y otras criaturas puede encontrarse en la estructura de la voluntad de una persona. Los seres humanos no son los únicos que tienen deseos y motivaciones ni los únicos que pueden elegir. Comparten estas cosas con los miembros de algunas otras especies, algunos de los cuales incluso parecen deliberar y tomar decisiones basadas en un pensamiento previo. No obstante, parece ser peculiarmente característico de los seres humanos el que puedan formar lo que denominaré “deseos de segundo orden” o “deseos del segundo orden”.

Además de querer, elegir y ser inducidos a *hacer* esto o aquello, es posible que los hombres también quieran tener (o no) ciertos deseos y motivaciones. Son capaces de querer ser diferentes, en sus preferencias y en sus propósitos, de lo que son. Muchos animales parecen tener la capacidad de lo que denominaré “deseos de primer orden” o “deseos del primer orden”, que simplemente son deseos de hacer o no una cosa u otra. Sin embargo, ningún animal, salvo el hombre, parece tener la capacidad de realizar la autoevaluación reflexiva que se manifiesta en la formación de los deseos de segundo orden.<sup>2</sup>

<sup>2</sup> En pos de la simplicidad, sólo me referiré a lo que alguien quiere o desea, sin tener en cuenta los fenómenos relacionados, como las elecciones y las decisiones. Propongo usar los verbos “querer” (*to want*) y “desear” (*to desire*) en forma indistinta, aunque de ninguna manera sean sinónimos perfectos. La razón de mi renuncia a los matices establecidos de estas palabras surge del hecho de que el verbo “querer”, que es más apropiado para mi objetivo en cuanto a su significado, no se presta con tanta facilidad a la formación de sustantivos como sucede con el verbo “desear”. En inglés, quizá sea aceptable, aunque poco elegante, hablar en plural de *the wants* (los “quereres”) de alguien. Sin embargo, hablar en singular de *the want* (el “querer”) de una persona resultaría una abominación.

I

El concepto que designa el verbo “querer” es en extremo esquivo. Una afirmación tal como “A quiere hacer X” –tomada en forma aislada, separada de un contexto que sirva para ampliar o especificar su significado– transmite extraordinariamente escasa información. Una afirmación como ésa puede ser coherente, por ejemplo, con cada una de las siguientes: a) la posibilidad de hacer X no provoca en A una sensación o respuesta emocional susceptible de introspección; b) A no es consciente de que quiere hacer X; c) A cree que no quiere hacer X; d) A quiere abstenerse de hacer X; e) A quiere hacer Y y cree que es imposible que pueda hacer tanto Y como X; f) A “en realidad” no quiere hacer X; g) A preferiría morir antes que hacer X; y así sucesivamente. Por tanto, no basta con formular la distinción entre los deseos de primer orden y los de segundo orden, como he hecho yo, sugiriendo simplemente que alguien tiene un deseo de primer orden cuando quiere hacer o no hacer tal cosa y que tiene un deseo de segundo orden cuando quiere tener o no tener cierto deseo del primer orden.

Según las interpretaré, afirmaciones tales como “A quiere hacer X” abarcan una gama de posibilidades bastante amplia.<sup>3</sup> Es posible que sean verdaderas incluso cuando son verdaderas afirmaciones como las identificadas con las letras (a) a (g): cuando A no es consciente de ningún sentimiento respecto de hacer X, cuando no sabe que quiere hacer X, cuando se engaña acerca de lo que quiere y cree falsamente que no quiere hacer X, cuando también tiene otros deseos que están en conflicto con su deseo de hacer X, o cuando es ambivalente. Los deseos en cuestión pueden ser conscientes o inconscientes, no es necesario que sean unívocos, y A puede estar errado acerca de ellos. Sin embargo, existe otra fuente de incertidumbre respecto de afirmaciones que identifican los de-

<sup>3</sup> Lo que expongo en este párrafo se aplica no sólo a los casos en que “hacer X” se refiere a una acción o a una inacción posibles. También se aplica a los casos en que “hacer X” se refiere a un deseo de primer orden y en los que la afirmación “A quiere hacer X” es, por tanto, una versión abreviada de una afirmación –“A quiere querer hacer X”– que identifica un deseo del segundo orden.

seos de alguien y, en este punto, es importante para mi propósito que sea menos permisivo.

Consideremos, en primera instancia, aquellas afirmaciones de la forma “A quiere hacer X” que identifican deseos de primer orden, es decir, afirmaciones en las que el término “hacer X” se refiere a una acción. Una afirmación de este tipo no indica, por sí misma, la fuerza relativa del deseo de A de hacer X. No deja en claro si es posible que este deseo desempeñe un papel decisivo en lo que A en realidad hace o intenta hacer. Por ello, se podría decir acertadamente que A quiere hacer X aun cuando su deseo de hacer X sea sólo uno entre sus deseos y aun cuando éste diste mucho de ser el principal. Por consiguiente, puede ser verdad que A quiera hacer X aunque en realidad prefiera hacer otra cosa; y puede ser verdad que quiera hacer X a pesar de que, cuando actúe, no sea el deseo de hacer X lo que lo motiva a hacer lo que hace. Por otro lado, alguien que afirma que A quiere hacer X podría querer expresar que es este deseo el que está motivando o induciendo a A a hacer lo que realmente está haciendo o que A, de hecho, será inducido por este deseo (a menos que cambie de parecer) cuando actúe.

Dado el uso especial de *voluntad* que propongo adoptar, sólo cuando la afirmación se emplea de la segunda manera identifica la voluntad de A. Identificar la voluntad de un agente significa ya identificar el deseo (o los deseos) que lo inducen a realizar alguna acción que lleva a cabo, ya identificar el deseo (o los deseos) que lo inducirán o lo inducirían cuando actúe o si actúa. Por tanto, la voluntad de un agente coincide con uno o más de sus deseos de primer orden. Pero la noción de voluntad, según la estoy empleando, no es coextensiva con la noción de deseos de primer orden. No es la noción de algo que simplemente inclina a un agente, en cierto grado, a actuar de cierta manera. Es, más bien, la noción de un deseo *efectivo*, tal que induce (o inducirá o induciría) a una persona a llevar a cabo la acción. Por tanto, la noción de voluntad no es coextensiva con la noción de lo que un agente tiene intenciones de hacer, ya que, aunque alguien pueda tener una intención firme de hacer X, tiene la posibilidad de hacer otra cosa en vez de hacer X porque, a pesar de su intención, su deseo de hacer X resulta ser más débil o menos efectivo que otro en conflicto con el primero.

Ahora consideremos las afirmaciones de la forma “A quiere hacer X” que identifican deseos de segundo orden, es decir, afirmaciones en las que el término “hacer X” se refiere a un deseo del primer orden. Hay también dos tipos de situaciones en que podría ser verdad que A quiere querer hacer X. En primer lugar, podría ser verdad que A quiere tener el deseo de hacer X, a pesar de tener un deseo unívoco, sin ningún tipo de conflicto ni ambivalencia, de abstenerse de hacer X. En otras palabras, alguien podría tener cierto deseo y, sin embargo, querer unívocamente que ese deseo no se satisfaga.

Supongamos que un médico especialista en psicoterapia con drogadictos cree que su capacidad de ayudar a sus pacientes mejoraría si pudiera comprender mejor qué significa para ellos desear la droga a la que son adictos. Supongamos que esto lo lleva a querer desear la droga. Si lo que quiere es sentir un deseo genuino, no sólo se trata, entonces, de experimentar las sensaciones que, por lo general, tienen los adictos cuando son dominados por sus deseos de consumir la droga. Lo que el médico quiere, en la medida en que quiere tener el deseo, es verse inclinado o inducido, en cierta medida, a consumirla.

Sin embargo, es totalmente posible que, aunque quiera sentirse inducido por el deseo de consumir la droga, no quiera que su deseo sea efectivo. Es posible que no quiera que el deseo lo induzca a la acción. No es necesario que esté interesado en descubrir qué se siente al consumir la droga. Y, en la medida en que ahora sólo quiere *querer* consumirla y no *consumirla*, no hay nada en lo que quiere en este momento que pueda ser satisfecho con la droga en sí misma. De hecho, es posible que ahora tenga un deseo totalmente unívoco de *no* consumir la droga; y es posible que disponga, prudentemente, que le sea imposible satisfacer el deseo que tendría si su deseo de querer la droga fuera satisfecho con el tiempo.

Por tanto, sería incorrecto inferir, a partir del hecho de que el médico ahora quiere desear consumir la droga, que ya tiene deseos de consumirla. Su deseo –de segundo orden– de ser inducido a consumir la droga no implica que sienta el deseo –de primer orden– de consumirla. Si se le fuera a administrar la droga, ello podría no satisfacer ningún deseo implícito en su deseo de querer

consumirla. Si bien quiere querer consumir la droga, es posible que *no* desee consumirla; podría ser que *todo* lo que quiere es probar el deseo de consumirla. Es decir, su deseo de tener cierto deseo que no tiene puede no constituir un deseo de que su voluntad sea en absoluto diferente de lo que es.

Alguien que, sólo de esta forma trunca, quiere querer hacer X se ubica en el margen del preciosismo, y el hecho de que quiere querer hacer X no es pertinente para la identificación de su voluntad. Sin embargo, hay un segundo tipo de situación que puede describirse como “A quiere querer hacer X”; y cuando se emplea la afirmación para describir una situación de este segundo tipo, en este caso sí se refiere a lo que A quiere que sea su voluntad. En dichos casos, la afirmación significa que A quiere que el deseo de hacer X sea el deseo que lo induzca a actuar efectivamente. No es sólo que quiere que el deseo de hacer X se encuentre entre los deseos por los cuales, en una u otra medida, es inducido o se siente inclinado a actuar. Quiere que este deseo sea efectivo, es decir, que proporcione la motivación de lo que realmente hace. Ahora bien, cuando la afirmación de que A quiere querer hacer X se emplea de esta manera, implica que A ya tiene el deseo de hacer X. No podría ser verdad a la vez que A quiere que el deseo de hacer X lo induzca a la acción y que no quiere hacer X. Sólo si realmente quiere hacer X podrá querer en forma coherente que el deseo de hacer X no sea sólo uno de sus deseos, sino que sea más decididamente su voluntad.<sup>4</sup>

Supongamos que un hombre quiere ser motivado en lo que hace por el deseo de concentrarse en su trabajo. Si esta suposición es correcta, necesariamente es verdad que ya quiere concentrarse en su

<sup>4</sup> No está tan claro que la relación de implicación aquí descrita se mantenga en ciertos tipos de casos que, pienso, podrían ser considerados con imparcialidad no estándar, donde la diferencia esencial entre los casos estándar y los no estándar reside en el tipo de descripción mediante la cual se identifica el deseo de primer orden en cuestión. Así, supongamos que A admira a B en forma tan exagerada que, incluso aunque no sepa qué quiere hacer B, A quiere ser inducido en forma efectiva por cualquier deseo que induzca en forma efectiva a B; en otras palabras, sin saber cuál es la voluntad de B, A quiere que su propia voluntad sea la misma. Por supuesto, no se desprende de ello que A ya tenga, entre sus deseos, un deseo como el que constituye la voluntad de B. No proseguiré aquí con preguntas acerca de si existen contraejemplos genuinos para la afirmación hecha en el texto o si, en caso de que existieran, esa afirmación debería modificarse.

trabajo. Este deseo es nuevo entre sus deseos. Sin embargo, la cuestión de si se cumple o no su deseo de segundo orden no depende sólo de si el deseo que quiere es uno de sus deseos. Depende de si este deseo es, tal como él quiere que sea, su deseo efectivo o voluntad. Si, a la hora de la verdad, es su deseo de concentrarse en su trabajo lo que lo induce a hacer lo que hace, entonces lo que quiere en ese momento es sin duda (en el sentido pertinente) lo que quiere querer. Por otra parte, si el deseo que lo induce realmente cuando actúa es algún otro, entonces lo que quiere en ese momento no es (en el sentido pertinente) lo que quiere querer. Ello será así aunque el deseo de concentrarse en su trabajo continúe estando entre sus deseos.

## II

Alguien tiene un deseo del segundo orden ya sea cuando simplemente quiere tener cierto deseo, ya cuando quiere que cierto deseo sea su voluntad. En situaciones de este último tipo, denominaré sus deseos de segundo orden “voliciones de segundo orden” o “voliciones del segundo orden”. Ahora bien, considero que es esencial tener voliciones de segundo orden y no deseos de segundo orden en general para ser una persona. Desde el punto de vista lógico es posible, aunque improbable, que exista un agente con deseos de segundo orden, pero sin voliciones del segundo orden. Dicha criatura, en mi opinión, no sería una persona. Emplearé el término *inconsciente*\* para referirme a agentes que tienen deseos de primer orden, pero que no son personas, porque, tengan o no deseos del segundo orden, no tienen voliciones de segundo orden.<sup>5</sup>

\* Del inglés *wanton*. En español, no se cuenta con un equivalente exacto de *wanton*, ya que en inglés esta palabra también incluye los distintos matices de “licencioso”, “libertino”, “desenfrenado”, “displicente”, “indisciplinado”e, incluso, “caprichoso”.  
[N. de T.]

<sup>5</sup> Las criaturas con deseos de segundo orden, pero sin voliciones de segundo orden difieren en gran medida de los animales brutos y, para algunos fines, sería conveniente considerarlas personas. Mi propuesta, que rechaza la denominación “persona” para ellos, es, por tanto, algo arbitraria. En gran parte, la adopto porque

La característica esencial de un agente inconsciente es que no le importa su voluntad. Sus deseos lo inducen a hacer ciertas cosas, sin que sea verdad ni que quiere ser inducido por esos deseos ni que prefiere ser inducido por otros deseos. La clase de agentes inconscientes incluye a todos los animales no humanos que tienen deseos y a todos los niños muy pequeños. Quizá también incluya a algunos seres humanos adultos. En todo caso, los seres humanos adultos pueden ser más o menos inconscientes; es posible que actúen de manera inconsciente en respuesta a deseos de primer orden respecto de los cuales no tienen voliciones del segundo orden, con mayor o menor frecuencia.

El hecho de que un agente inconsciente no tenga voliciones de segundo orden no significa que cada uno de sus deseos de primer orden se traduzca en forma irreflexiva y de inmediato en acción. Quizá no tenga la oportunidad de actuar de acuerdo con algunos de sus deseos. Más aun, la traducción de sus deseos en acción puede ser demorada o impedida ya sea por deseos conflictivos del primer orden, ya por la deliberación. Esto se debe a que es posible que un agente inconsciente posea y emplee facultades racionales de una jerarquía alta. No hay nada en el concepto de agente inconsciente que implique que éste no pueda razonar o deliberar acerca de cómo hacer lo que quiere hacer. Lo que distingue al agente inconsciente racional de otros agentes racionales es que no le preocupa la conveniencia de sus deseos. Hace caso omiso de la pregunta acerca de cuál ha de ser su voluntad. No sólo sigue cualquier proceder que se vea más fuertemente inclinado a seguir, sino que no le importa cuál de sus inclinaciones es la más fuerte.

Por tanto, una criatura racional, que reflexiona acerca de la conveniencia de sus deseos de seguir un curso de acción u otro, puede ser, sin embargo, inconsciente. Al sostener que la esencia de ser una persona reside no en la razón sino en la voluntad, estoy lejos de sugerir que una criatura sin razón puede ser una persona. Ello se

facilita la formulación de algunas de las observaciones que quiero hacer. De aquí en adelante, cada vez que considere afirmaciones tales como “A quiere querer hacer X”, tendré en cuenta afirmaciones que identifiquen voliciones de segundo orden y no afirmaciones que identifiquen deseos de segundo orden que no son voliciones de segundo orden.

debe a que sólo en virtud de sus capacidades racionales una persona es capaz de volverse críticamente consciente de su propia voluntad y de formar voliciones del segundo orden. La estructura de la voluntad de una persona presupone, en consecuencia, que es un ser racional.

La distinción entre una persona y un agente inconsciente puede ilustrarse mediante la diferencia entre dos drogadictos. Supongamos que la condición fisiológica responsable de la adicción es la misma en ambos hombres, y que ambos sucumben de manera inevitable a sus deseos periódicos de consumir la droga a la cual son adictos. Uno de los adictos detesta su adicción y siempre lucha desesperadamente, aunque en vano, contra su embestida. Prueba todo lo que piensa que podría permitirle superar sus deseos de consumir la droga. Sin embargo, estos deseos son demasiado poderosos para él y, al final, invariablemente lo conquistan. Se trata de un adicto contra su voluntad, indefenso ante sus propios deseos.

El adicto contra su voluntad tiene deseos conflictivos de primer orden: quiere consumir la droga y también quiere abstenerse de hacerlo. Sin embargo, además de estos deseos de primer orden, tiene una volición del segundo orden. No es neutral respecto del conflicto entre su deseo de consumir la droga y su deseo de abstenerse de hacerlo. Este último deseo, y no el primero, es el que él quiere que constituya su voluntad; el último deseo y no el primero es el que él quiere que sea efectivo y le brinde el propósito que intentará concretar en lo que haga realmente.

El otro adicto es un agente inconsciente. Sus acciones reflejan la economía de sus deseos de primer orden, sin que le preocupe si los deseos que lo inducen a actuar son los deseos por los cuales quiere ser inducido a actuar. Si tiene problemas para obtener la droga o para administrársela, sus respuestas a su necesidad de consumirla podrían requerir deliberación. Sin embargo, nunca se le ocurre considerar si quiere que las relaciones entre sus deseos tengan como resultado que él tenga la voluntad que tiene. El adicto inconsciente puede ser un animal y, por tanto, ser incapaz de preocuparse por su voluntad. En cualquier caso, en lo que concierne a la inconciencia de su falta de preocupación, no es diferente de un animal.

El segundo de estos adictos puede sufrir un conflicto de primer orden similar al conflicto del primer orden que sufre el primer adicto. Sea humano o no, es posible que el agente inconsciente (quizá debido al condicionamiento) quiera tanto consumir la droga como abstenerse de consumirla. A diferencia del adicto contra su voluntad, sin embargo, no prefiere que uno de sus deseos conflictivos prevalezca sobre el otro; no prefiere que un deseo de primer orden, y no otro, constituya su voluntad. Sería engañoso decir que es neutral en cuanto al conflicto entre sus deseos, puesto que ello sugeriría que los considera igualmente aceptables. Y ya que no tiene otra identidad que sus deseos de primer orden, no es cierto ni que prefiera uno antes que el otro ni que prefiera no tomar partido.

Para el adicto contra su voluntad, que es una persona, importa mucho cuál de sus deseos conflictivos de primer orden gana. Ambos deseos son suyos, por cierto; y si finalmente consume la droga o finalmente logra abstenerse de consumirla, actúa para satisfacer lo que es, en sentido literal, su propio deseo. En cualquiera de los casos, hace algo que él mismo quiere hacer, y no lo hace debido a cierta influencia externa cuyo objetivo coincide por casualidad con el propio, sino por su deseo de hacerlo. No obstante, el adicto contra su voluntad se identifica a sí mismo mediante la formación de una volición de segundo orden con uno de sus deseos conflictivos de primer orden y no con el otro. Hace que uno de ellos sea más auténticamente suyo y, al hacerlo, se distancia del otro. En virtud de esta identificación y este distanciamiento, que se logran mediante la formación de una volición de segundo orden, el adicto contra su voluntad puede, en forma significativa, hacer las declaraciones –desconcertantes desde el punto de vista analítico– de que la fuerza que lo induce a consumir la droga es una fuerza distinta de la propia, y que esta fuerza lo induce a consumirla no por su propio libre albedrío, sino, por el contrario, contra su voluntad.

Al adicto inconsciente no le puede importar o no le importa cuál de sus deseos conflictivos de primer orden gana. Su falta de preocupación no se debe a su incapacidad de encontrar una base convincente para sus preferencias. Se debe o bien a su falta de capacidad de reflexión o bien a su indiferencia mecánica frente a la

tarea de evaluar sus propios deseos y motivos.<sup>6</sup> Hay una sola cuestión en la lucha a la cual puede conducir este conflicto de primer orden: cuál de sus deseos conflictivos es más fuerte. Debido a que es inducido por ambos deseos, no sentirá plena satisfacción por lo que haga independientemente de cuál de ellos sea efectivo. Pero *para él* es lo mismo si se imponen sus anhelos o su aversión. A él no le interesa el conflicto que existe entre ellos y, por tanto, a diferencia del adicto contra su voluntad, no puede ni ganar ni perder la lucha en la que está empeñado. Cuando una *persona* actúa, el deseo por el cual es inducido a la acción es o bien la voluntad que quiere o bien una voluntad que no quiere tener. Cuando un *agente inconsciente* actúa, no se trata de ninguna de las dos.

### III

Existe una relación muy estrecha entre la capacidad de formar voliciones de segundo orden y otra capacidad que es esencial para las personas: una que a menudo ha sido considerada una marca distintiva de la condición humana. El hecho de que una persona sea capaz tanto de disfrutar como de carecer de la libertad de la voluntad se debe sólo a que tiene voliciones del segundo orden. El concepto de persona, por consiguiente, no es sólo el concepto de un tipo de entidad que tiene tanto deseos de primer orden como voliciones del segundo orden. También se lo puede interpretar como el concepto de un tipo de entidad para la que la libertad de su voluntad podría representar un problema. Este concepto ex-

<sup>6</sup> Cuando digo que la evaluación de sus propios deseos y de sus motivos es característica de una persona, no quiero sugerir que las voliciones de segundo orden de una persona manifiesten necesariamente una postura *moral* de su parte respecto de sus deseos de primer orden. Es posible que la persona no evalúe sus deseos de primer orden desde el punto de vista de la moralidad. Más aun, es posible que una persona actúe en forma caprichosa e irresponsable al formar sus voliciones de segundo orden y no considere con seriedad qué está en juego. Las voliciones de segundo orden expresan evaluaciones sólo en el sentido de que son preferencias. No hay una restricción esencial en cuanto al tipo de fundamento, si es que lo hay, sobre el cual se forman.

cluye a todos los agentes inconscientes, tanto infrahumanos como humanos, puesto que no satisfacen una condición esencial para gozar de la libertad de la voluntad. Y excluye a esos seres sobrehumanos, si los hay, cuya voluntad es necesariamente libre.

De hecho, ¿qué tipo de libertad es la libertad de la voluntad? Esta pregunta requiere una identificación del área especial de la experiencia humana con la que el concepto de libertad de la voluntad, a diferencia de los conceptos de otros tipos de libertad, está particularmente vinculado. Al abordarla, mi objetivo será, en primer lugar, localizar el problema con el que se enfrenta una persona de manera inmediata cuando se interesa por la libertad de su voluntad.

Según una tradición filosófica conocida, ser libre es, fundamentalmente, cuestión de hacer lo que uno quiere hacer. Ahora bien, la noción de un agente que hace lo que quiere hacer no es, de ninguna manera, totalmente clara: tanto el hacer como el querer, como también la adecuada relación entre ellos, tienen que dilucidarse. Pero a pesar de que esta noción requiere una mayor exactitud y su formulación tiene que refinarse, creo que al menos captura parte de lo que está implícito en la idea de un agente que *actúa* libremente. No obstante, pierde por completo el contenido peculiar de la idea —bastante diferente— de un agente cuya *voluntad* es libre.

No suponemos que los animales gozan de libertad de la voluntad, aunque reconocemos que un animal puede ser libre para correr en la dirección que quiera. Por consiguiente, tener la libertad de hacer lo que uno quiere hacer no es una condición suficiente para tener libre albedrío. Tampoco es una condición necesaria, puesto que privar a alguien de su libertad de acción no significa necesariamente debilitar la libertad de su voluntad. Cuando un agente es consciente de que hay ciertas cosas que no es libre de hacer, ello, sin dudas, afecta sus deseos y limita el rango de las elecciones que puede hacer. Pero supongamos que alguien, sin saberlo, ha perdido o ha sido privado de su libertad de acción. Aunque ya no tiene la libertad de hacer lo que quiere hacer, es probable que su voluntad siga siendo tan libre como lo era anteriormente. A pesar del hecho de que no es libre para traducir sus deseos en acciones o para actuar según lo que determina su voluntad, aún podrá formar

esos deseos y tomar esas determinaciones tan libremente como si su libertad de acción no hubiera sido afectada.

Cuando preguntamos si la voluntad de una persona es libre, no estamos preguntando si tiene la posibilidad de traducir sus deseos de primer orden en acciones. Eso sería preguntar si es libre para hacer lo que le plazca. La pregunta por la libertad de su voluntad no tiene que ver con la relación entre lo que hace y lo que quiere hacer. Más bien, tiene que ver con sus deseos en sí mismos. Pero, ¿qué sucede con ellos?

Me parece natural y útil interpretar la pregunta de si la voluntad de una persona es libre en estrecha analogía con la pregunta de si un agente goza de libertad de acción. Ahora bien, la libertad de acción es (aproximadamente, al menos) la libertad de hacer lo que uno quiere hacer. De manera similar, entonces, la afirmación de que una persona goza de libertad de la voluntad significa (también aproximadamente) que es libre de querer lo que quiere querer. Para ser más exacto, significa que es libre de desear lo que quiera desear, o de tener el deseo que quiera. Así como la cuestión de la libertad de acción de un agente está relacionada con el hecho de si se trata de la acción que quiere realizar, la cuestión de la libertad de su voluntad está relacionada con el hecho de si es la voluntad que quiere tener.

Por tanto, una persona ejercita la libertad de la voluntad asegurando la conformidad de su voluntad con sus voliciones de segundo orden. Y la discrepancia entre su voluntad y sus voliciones de segundo orden o el hecho de saber que su coincidencia no es obra suya, sino sólo una casualidad, hacen que la persona que no tiene esta libertad sienta su carencia. La voluntad del adicto contra su voluntad no es libre. Esto se ve en el hecho de que no es la voluntad que quiere. También es verdad, aunque de manera diferente, que la voluntad del adicto inconsciente no es libre. El adicto inconsciente no tiene ni la voluntad que quiere ni una voluntad diferente de la voluntad que quiere. Dado que no tiene voliciones del segundo orden, la libertad de su voluntad no puede ser un problema para él. Desde siempre, por así decirlo, carece de ella.

Por lo general, la gente es mucho más complicada de lo que podría sugerir mi esbozo sobre la estructura de la voluntad de una

persona. Hay tantas oportunidades para las ambivalencias, el conflicto y el autoengaño respecto de los deseos del segundo orden, por ejemplo, como respecto de los deseos de primer orden. Si hay un conflicto no resuelto entre los deseos de segundo orden de alguien, esta persona está en peligro de no tener volición de segundo orden, puesto que, a menos que se resuelva este conflicto, no tiene preferencias respecto de cuál de sus deseos de primer orden ha de ser su voluntad. Esta condición, si es tan seria como para impedirle identificarse, en forma lo suficientemente decisiva, con *alguno* de sus deseos conflictivos de primer orden, lo destruye como persona. Esto se debe a que tiende a paralizar su voluntad y a impedir que actúe o tiende a distanciarlo de su voluntad de manera que ésta opera sin su participación. En ambos casos, tal como el adicto contra su voluntad aunque en forma diferente, se convierte en un espectador indefenso de las fuerzas que lo inducen.

Otra cuestión compleja es que la persona puede tener, en especial si sus deseos de segundo orden están en conflicto, deseos y voliciones de un orden superior al segundo. No existe un límite teórico para la extensión de la serie de deseos de órdenes más y más altos; no existe nada, salvo el sentido común y, quizá, una fatiga salvadora que impide que un individuo se niegue, en forma obsesiva, a identificarse con alguno de sus deseos hasta que forma un deseo del orden superior siguiente. La tendencia a generar una serie tal de actos de formación de deseos, que constituiría un caso de humanización sin freno, también conduce a la destrucción de una persona.

No obstante, es posible poner fin a tal serie de actos sin cortarla en forma arbitraria. Cuando una persona se identifica *decididamente* con uno de sus deseos de primer orden, este compromiso “resuena” a través de la serie potencialmente infinita de órdenes superiores. Consideremos una persona que, sin reserva ni conflictos, quiere ser motivada por el deseo de concentrarse en su trabajo. El hecho de que su volición de segundo orden de ser inducida por este deseo es decisiva significa que no hay lugar para preguntas acerca de la pertinencia de deseos o voliciones de órdenes superiores. Supongamos que se le pregunta a la persona si quiere querer querer concentrarse en su trabajo. Ella bien puede insistir en

que no se plantea la pregunta sobre un deseo de tercer orden. Sería erróneo sostener que, debido a que no ha considerado si quiere la volición de segundo orden que ha formado, es indiferente a la pregunta de si quiere que su voluntad concuerde con esta volición o con alguna otra. La firmeza del compromiso que ha contraído significa que ha decidido que no queda por formular ninguna otra pregunta acerca de su volición de segundo orden, en ningún orden superior. Es relativamente poco importante la cuestión de si lo explicamos diciendo que este compromiso genera, en forma implícita, una serie infinita de deseos de confirmación de órdenes superiores o que el compromiso es equivalente a una disolución de la importancia de todas las preguntas relacionadas con órdenes superiores de deseos.

Ejemplos como el del adicto contra su voluntad podrían sugerir que las voliciones del segundo orden, o de órdenes superiores, deben formarse de manera deliberada y que, característicamente, una persona lucha para asegurar que se satisfagan. Pero la conformidad de la voluntad de una persona a sus voliciones de orden superior puede ser mucho más irreflexiva y espontánea. Algunas personas son inducidas en forma natural por la bondad cuando quieren ser buenas, y por la maldad cuando quieren ser malas, sin ninguna reflexión previa explícita y sin la necesidad de un autocontrol firme. Otras son inducidas por la maldad cuando quieren ser buenas y por la bondad cuando tienen intenciones de ser malas, también sin ninguna reflexión previa y sin una resistencia activa a estas violaciones de sus deseos de orden superior. Gozar de la libertad les es fácil a algunos. Otros deben luchar por lograrlo.

#### IV

Mi teoría respecto de la libertad de la voluntad explica fácilmente nuestra reticencia a aceptar que los miembros de cualquier especie inferior a la nuestra puedan disfrutar de esta libertad. Asimismo, cumple otra condición que cualquier teoría de este tipo debe satisfacer, evidenciando por qué debe considerarse deseable

la libertad de la voluntad. Disfrutar de libre albedrío supone la satisfacción de ciertos deseos —deseos del segundo orden o de órdenes superiores—, mientras que su ausencia significa su frustración. Las satisfacciones en juego son aquellas que se le confieren a una persona de quien se puede decir que tiene voluntad propia. Las frustraciones correspondientes son aquellas que sufre una persona de quien se puede decir que está alejada de sí misma, o que se considera un espectador indefenso o pasivo frente a las fuerzas que lo inducen.

Una persona que es libre de hacer lo que quiere hacer puede, a pesar de ello, estar privada de la voluntad que quiere. Supongamos, sin embargo, que goza tanto de libertad de acción como de libertad de la voluntad. Entonces, no sólo es libre de hacer lo que quiere hacer; también es libre de querer lo que quiere querer. Me parece que, en ese caso, tiene toda la libertad que es posible desear o concebir. Hay otras cosas buenas en la vida, y es posible que no posea algunas de ellas. Pero no carece de nada en materia de libertad.

No es evidente en absoluto que algunas otras teorías sobre la libertad de la voluntad satisfagan estas condiciones elementales pero básicas: que sea comprensible por qué deseamos esta libertad y por qué rechazamos atribuírsela a los animales. Consideremos, por ejemplo, la curiosa versión de Roderick Chisholm de la doctrina de que la libertad humana implica una ausencia de determinación causal.<sup>7</sup> Cuando una persona lleva a cabo una acción libre, según Chisholm, se trata de un milagro. El movimiento de la mano de una persona, cuando ésta la mueve, es el resultado de una serie de causas físicas; pero algún acontecimiento de esta serie, “y presumiblemente uno de los que se produjeron en el cerebro, fue causado por el agente y por ningún otro acontecimiento” (p. 18). Un agente libre tiene, por tanto, “una prerrogativa que algunos sólo le atribuirían a Dios: cada uno de nosotros, cuando actuamos, es la principal fuerza motriz impasible” (p. 23).

Esta explicación no ofrece ningún fundamento para dudar de que los animales de especies subhumanas gocen de la libertad que

7 R. Chisholm, “Freedom and Action”, en K. Lehrer (ed.), *Freedom and determinism*, Nueva York, Random House, 1966, pp. 11-44.

define. Nada de lo que dice Chisholm nos permite considerar menos probable el carácter milagroso de la acción de un conejo al mover la pata que el de la acción de un hombre al mover la mano. Pero, en todo caso, ¿por qué le debería *importar* a alguien la posibilidad de interrumpir el orden natural de las causas de la manera que describe Chisholm? Chisholm no da razones para creer que existe una diferencia discernible entre la experiencia de un hombre que inicia milagrosamente una serie de causas cuando mueve la mano y un hombre que mueve la mano sin violar la secuencia causal normal. No parecería haber fundamentos concretos para preferir estar envuelto en una situación más que en la otra.<sup>8</sup>

Por lo general se supone que, además de cumplir con las dos condiciones que he mencionado, una teoría satisfactoria de la libertad de la voluntad necesariamente proporciona un análisis de una de las condiciones de la responsabilidad moral. El abordaje reciente más común para el problema de cómo comprender la libertad de la voluntad ha sido, de hecho, indagar qué implica la suposición de que alguien es moralmente responsable de lo que hizo. No obstante, en mi opinión, la relación entre responsabilidad moral y la libertad de la voluntad ha sido, por lo general, mal entendida. No es verdad que una persona sea moralmente responsable de lo que hizo sólo si su voluntad era libre cuando lo hizo. Es posible que sea moralmente responsable de haberlo hecho incluso si su voluntad no era en absoluto libre.

La voluntad de una persona es libre sólo si esa persona es libre para tener la voluntad que quiere. Esto significa que, respecto de cualquiera de sus deseos de primer orden, es libre ya sea de hacer que ese deseo sea su voluntad, ya de hacer que algún otro deseo de primer orden sea su voluntad. Sea cual fuere su voluntad, entonces, la voluntad de la persona cuya voluntad es libre podría haber sido otra; podría haber hecho otra cosa diferente que constituir su voluntad tal como lo hizo. La cuestión de cómo entender “podría

haber hecho otra cosa” en contextos como éste resulta polémica. Sin embargo, aunque esta cuestión es importante para la teoría de la libertad, no tiene relación con la teoría de la responsabilidad moral, ya que la suposición de que una persona es moralmente responsable de lo que hizo no implica que la persona estuviera en una posición de tener la voluntad que quería.

Lo que esta suposición sí implica es que la persona hizo lo que hizo libremente, o que lo hizo por su propio libre albedrío. No obstante, es un error creer que alguien actúa libremente sólo cuando es libre de hacer lo que quiere o que actúa según su libre albedrío si su voluntad es libre. Supongamos que una persona hizo lo que quería hacer, que lo hizo porque quería hacerlo, y que la voluntad por la cual fue inducida cuando lo hizo era su voluntad, porque era la voluntad que quería. Entonces, lo hizo libremente y por su propio libre albedrío. Aun en el supuesto de que hubiera podido hacer otra cosa, no habría querido hacer otra cosa; y aun en el supuesto de que hubiera podido tener una voluntad diferente, no habría querido que su voluntad difiriera de la que era. Además, debido a que la voluntad que la indujo cuando actuó era su voluntad porque ella quería que así lo fuera, no puede alegar que fue obligada a tener esa voluntad ni que asistió pasivamente a su constitución. En estas condiciones, es bastante irrelevante para la evaluación de su responsabilidad moral preguntar si realmente tenía a su disposición las alternativas que descartó.<sup>9</sup>

A manera de ejemplo, consideremos un tercer tipo de adicto. Supongamos que su adicción tiene la misma base fisiológica y la misma embestida irresistible que las adicciones de los adictos contra su voluntad y los adictos inconscientes, pero que está totalmente encantado con su condición. Es un adicto por voluntad propia, que no querría que las cosas fueran distintas. Si la fuerza de su adicción en cierta forma se debilitara, haría todo lo que estuviera a su alcance para reavivarla; si su deseo por la droga comenzara a mermar, tomaría medidas para renovar su intensidad.

<sup>8</sup> No estoy sugiriendo que la supuesta diferencia entre estas dos situaciones sea inverificable. Por el contrario, los fisiólogos bien podrían demostrar que las condiciones impuestas por Chisholm para definir una acción libre no se cumplen, estableciendo que no hay ningún acontecimiento mental pertinente para el cual no se pueda encontrar una causa física suficiente.

<sup>9</sup> Véase en “Posibilidades alternativas y responsabilidad moral”, capítulo 1 de este volumen, otro análisis de las consideraciones que ponen en duda el principio de que una persona es moralmente responsable de lo que hizo sólo si podría haber hecho otra cosa.

La voluntad del adicto por voluntad propia no es libre, puesto que su deseo de consumir la droga será efectivo independientemente de si quiere o no que este deseo constituya su voluntad. Pero cuando consume la droga, lo hace libremente y por propio libre albedrío. Estoy inclinado a interpretar su situación como sobredeterminada por su deseo de primer orden de consumir la droga. Este deseo es su deseo efectivo porque fisiológicamente es adicto. Sin embargo, también es su deseo efectivo porque él quiere que lo sea. Su voluntad está fuera de su control, pero, debido a su deseo de segundo orden de que su deseo de la droga sea efectivo, ha convertido a esta voluntad en propia. Debido a que su deseo de la droga es efectivo no sólo por su adicción, es posible que sea moralmente responsable de consumirla.

Mi idea de la libertad de la voluntad es, en apariencia, neutral respecto del problema del determinismo. Se puede pensar que el hecho de que una persona sea libre de querer lo que quiere querer esté causalmente determinado. Si esto es concebible, entonces el que una persona goce de libre albedrío también podría estar determinado causalmente. No hay más que una inocua apariencia de paradoja en la proposición que sostiene que está establecido, en forma inevitable y por fuerzas más allá de su control, que ciertas personas tienen libre albedrío y otras no. No hay incoherencia en la proposición de que una agencia, diferente de la propia de la persona, sea responsable (incluso *moralmente* responsable) del hecho de que esa persona disfrute o no de la libertad de la voluntad. Es posible que una persona sea moralmente responsable de lo que hace por su propio libre albedrío, y que alguna otra persona también sea moralmente responsable de que la primera lo haya hecho.<sup>10</sup>

<sup>10</sup> Existe una diferencia entre ser *completamente* responsable y ser *únicamente* responsable. Supongamos que el adicto por voluntad propia se ha hecho adicto por el trabajo deliberado y calculado de otro. Entonces, podría ser que tanto el adicto como esta otra persona fueran completamente responsables de que el adicto consuma la droga, mientras que ninguno de ellos es únicamente responsable de ello. El hecho de que existe una distinción entre la responsabilidad moral completa y la responsabilidad moral única es evidente en el siguiente ejemplo. Hay una luz que puede encenderse o apagarse activando uno de dos interruptores, y cada uno de estos interruptores es activado simultáneamente a la posición de "encendido" por dos personas diferentes que desconocen la existencia

Por otra parte, podemos imaginar que podría suceder por casualidad que una persona fuera libre de tener la voluntad que quiere. Si cabe la posibilidad de que esto suceda, podría ser cuestión de suerte que algunas personas gocen de la libertad de la voluntad, y que otras no lo hagan. Quizá también se pueda imaginar, como creen varios filósofos, que las situaciones acontecen de manera no casual o como resultado de una secuencia de causas naturales. Si en realidad se puede concebir que las situaciones pertinentes acontezcan de una manera diferente de las dos anteriores, también es posible que una persona, en esa tercera manera, llegue a gozar de la libertad de la voluntad.

de la otra. Ninguna de las personas es únicamente responsable de que la luz esté encendida, ni tampoco comparten la responsabilidad en el sentido de que cada una es parcialmente responsable; más bien, cada una de ellas es completamente responsable.