

(Pre-print)

LIBERTAD, DETERMINISMO Y RESPONSABILIDAD MORAL*

Carlos J. Moya. Universidad de Valencia (carlos.moya@uv.es)

El problema de las relaciones entre libertad y determinismo tiene una importante presencia en la historia de la filosofía, y sigue siendo objeto de una viva discusión en la actualidad. En el siglo XII, el poeta y pensador sufí Jalalu'ddin Rumi escribía: “Hay una disputa que continuará hasta que el género humano se levante de entre los muertos, entre los partidarios de la necesidad y los defensores del libre albedrío” (cf. Kane 2009, p. 26). Varios siglos después, Hume consideraba “la cuestión de la libertad y la necesidad [como] la más controvertida cuestión de la metafísica, la ciencia más controvertida” (Hume 1748/2004, p. 215). Un aspecto de esta cuestión lo constituye, como vio el propio Hume, la relación del determinismo con la responsabilidad moral, dado que la libertad es ampliamente considerada como un requisito de esta última. Y la controversia sigue siendo en la actualidad tan reñida como lo fue entonces. El presente trabajo está dedicado a este debate. En la primera sección analizaremos brevemente los conceptos de determinismo, libertad y responsabilidad moral y presentaremos las principales posiciones acerca de sus relaciones. En las secciones siguientes abordaremos distintos aspectos de esta problemática.

I. CONCEPTOS Y POSICIONES FUNDAMENTALES

1. *El determinismo*

De acuerdo con el determinismo (causal), lo que sucede en cada momento de la historia del universo es necesario que suceda (no puede no suceder) dado el estado anterior del universo y las leyes naturales. De modo más preciso, siguiendo a Peter Van Inwagen, el determinismo es la conjunción de dos tesis. La primera es: “Para cada instante, hay una proposición que expresa el estado del mundo en ese instante”. La segunda es: “Si p y q son proposiciones cualesquiera que expresan el estado del mundo en algunos instantes, la conjunción de p y de las leyes de la naturaleza implica q ” (Van

* Este trabajo se encuadra en el proyecto de investigación “Alcance y límites de la responsabilidad” (FFI2012-33470), otorgado por la Dirección General de Investigación, Ministerio de Economía y Competitividad. Agradezco a Josep Corbí, Alex Fontcuberta, Milagro Garrido, Tobies Grimaltos, Edgar Maragat, Eduardo Ortiz y Pablo Rychter sus observaciones y sugerencias sobre una versión anterior del presente trabajo.

Inwagen 1983, p. 65). Así, un ser que conociera por completo el estado del mundo en un momento dado y tuviera asimismo un conocimiento perfecto de las leyes de la naturaleza, podría saber, por razonamiento deductivo, cuál sería el estado del mundo en cualquier otro momento. Sólo un ser omnisciente podría llevar a cabo esta descomunal hazaña cognoscitiva; aunque el determinismo sea verdadero, dado que los seres humanos no somos omniscientes, no podremos llevarla a cabo, y muchos sucesos se nos presentarán como puramente casuales. Pero si el determinismo es verdadero, nada sucede por casualidad. La casualidad, el azar, son, desde esta perspectiva, conceptos meramente epistémicos, no metafísicos: lo casual, lo azaroso, son apariencias producidas por la limitación de nuestras capacidades cognoscitivas, no rasgos reales del mundo.

La tesis del determinismo causal es metafísica, en el doble sentido de que es una tesis sobre la naturaleza de la realidad y de que rebasa el alcance de la física y de las ciencias empíricas en general. Nadie sabe realmente si es verdadera, pero muchos creen o están convencidos de que lo es, aun cuando, debido a la presencia de teorías físicas como la mecánica cuántica, la tesis determinista no posee ya, entre los científicos y los filósofos, el aura de axioma o presupuesto obvio que pudo tener desde el desarrollo de la Nueva Ciencia, y de la mecánica newtoniana en particular, hasta principios del siglo XX. Sin embargo, a pesar de que su formulación es relativamente sencilla, su contenido dista mucho de ser claro. Algunos pensadores, como Peter Strawson, declararon no entender en qué consistía exactamente la tesis del determinismo¹. Y no es difícil comprender la perplejidad de Strawson, ya que las definiciones del determinismo que arriba hemos recogido incluyen conceptos de dudosa inteligibilidad, como el de una descripción completa del estado del universo en un instante dado. Pensemos en la enorme diversidad de entidades y de propiedades que pueblan el universo, y en los problemas que plantea su identidad y clasificación según los distintos tipos de discurso, desde el científico al ético o estético. Teniendo esto en cuenta, ¿podemos dar un contenido inteligible a la idea de una descripción completa del universo, incluso para un ser omnisciente?

Un modo de hacerlo podría ser limitar el alcance de la descripción concibiéndola como una descripción *física* completa del universo, una descripción del universo en términos de la física básica, convenientemente idealizada. La noción de una descripción física completa del universo en un instante determinado parece inteligible, aun cuando su

¹ Aun cuando ello no le impidió discutir el problema de sus relaciones con la libertad y la responsabilidad moral, asumiendo una comprensión intuitiva del determinismo (cf. Strawson 1962),

logro siga estando más allá de nuestras capacidades. Por otra parte, para que sea posible deducir, a partir de una descripción física completa del universo en un instante determinado, junto con las leyes físicas, la descripción física completa del universo en otro instante dado, es necesario que dichas leyes físicas sean estrictamente deterministas, no meramente probabilísticas. Supongamos, por mor del argumento, que lo son. Dadas estas condiciones, la tesis del determinismo, referida al mundo físico, parece al menos concebible y susceptible de verdad o falsedad.

Supongamos además que hay una relación robusta de superveniencia o dependencia metafísica entre las propiedades físicas y cualesquiera otras propiedades del mundo, de modo que, necesariamente, fijadas aquéllas, todas las demás propiedades quedan también fijadas. Entre éstas se hallarían, pues, propiedades psicológicas especialmente relevantes para las cuestiones relativas a la libertad y la responsabilidad moral, como deseos, creencias, propósitos, valores, decisiones, preferencias, etc. Dado este supuesto, el determinismo físico implicaría el determinismo con respecto a cualquier otro aspecto de la realidad, incluyendo especialmente el determinismo en relación con las decisiones y acciones humanas. Así, un ser omnisciente podría conocer, partiendo de una descripción física completa del mundo en un instante dado, junto con el conocimiento de las leyes físicas y de las relaciones de superveniencia entre las propiedades físicas y cualesquiera otras, cuáles serían los pensamientos, elecciones y acciones de cualquier ser humano en cualquier instante dado.

Como he indicado, no sabemos si la tesis del determinismo es verdadera o falsa; sin embargo, interpretada en los términos precedentes, es una tesis con un contenido inteligible y podría, así, ser verdadera. Es importante entonces reflexionar sobre las consecuencias que tendría su verdad (y su falsedad) sobre la libertad y la responsabilidad moral, como propiedades que, supuestamente, caracterizan a los seres humanos, o, más exactamente, a muchos seres humanos adultos. Para ello, sin embargo, hemos de preguntarnos en qué consisten dichas propiedades. Ocupémonos ahora, pues, de la libertad.

2. La libertad

La libertad puede entenderse, en una primera aproximación, como una capacidad o poder de determinados agentes para llevar a cabo ciertos actos con un grado apropiado

de control sobre los mismos². Y de los actos en cuestión decimos que son libres, o que han sido realizados libremente, cuando resultan del ejercicio de dicha capacidad, con el grado apropiado de control. Pero ¿qué aspectos o condiciones comprende la libertad, o, en otras palabras, el control apropiado en el ejercicio de la capacidad de actuar? Podemos tratar de responder a esta pregunta por vía negativa, considerando casos en los cuales rechazaríamos, intuitivamente, el carácter libre de la acción de que se trate.

En primer lugar, no consideramos libre una acción que el sujeto ha llevado a cabo sin darse cuenta o que, aun dándose cuenta, no llevó a cabo voluntariamente, o ambas cosas. Despertar al bebé que, sin yo saberlo, duerme en una habitación del piso del vecino al aumentar el volumen de mi televisor sería un ejemplo que satisface ambas condiciones. Es cierto que he despertado al bebé, pero no que lo haya hecho intencionalmente. Podemos resumir esta idea diciendo que no consideramos libre una acción no intencional. Así, ser intencional es una condición necesaria del carácter libre de una acción.

Que una acción sea intencional, sin embargo, no es suficiente para que sea libre. Pensemos en un empleado de un banco que entrega a un atracador, que le apunta con un revólver, el dinero de la caja. Hay un sentido claro en que el empleado quiere entregarle el dinero al atracador, con el fin de conservar su vida, y desde luego es consciente de que se lo entrega. La acción de entregar el dinero fue, pues, intencional por parte del empleado. Pero no diríamos que entregarle el dinero fue algo que hizo libremente. Y tampoco fue libre la decisión de hacerlo.

¿Qué más se requiere, además de ser intencional, para que una acción se lleve a cabo libremente? En el ejemplo anterior, aun cuando el empleado tuvo algún grado de control sobre su acción, ese grado no fue suficiente para que actuase libremente. En primer lugar, la alternativa de no entregar el dinero, aun cuando estrictamente no estaba excluida, tenía un coste previsible tan alto (la pérdida de su propia vida) que, en la práctica, no contaba como tal. El empleado habría podido decir, sin faltar a la verdad, que no tenía otra alternativa que entregar el dinero. No tenía control suficiente sobre su acción porque no tenía otra alternativa a su alcance. Así, que un agente tenga alternativas de decisión o acción que estén efectivamente a su alcance o bajo su control efectivo parece asimismo una condición para que la decisión que realmente toma y la acción que realmente lleva a cabo se puedan considerar libres. Este aspecto de la libertad es tan importante que se ha tendido a identificarlo con ella. Así, Thomas Pink escribe: “Por

² Esta última restricción excluye de la libertad, por ejemplo, la capacidad de digerir los alimentos que uno ingiere.

libertad entiendo la libertad de las posibilidades alternativas: la libertad de hacer cosas o no hacerlas o –como también puedo formularlo– el *control* sobre si hacemos esas cosas o no” (Pink 1996, p. 1). De acuerdo con esta concepción, ciertamente plausible y ampliamente aceptada, la libertad puede entenderse como la capacidad de elegir entre alternativas efectivamente accesibles al agente, sobre las cuales el agente tiene control efectivo, un control que está ausente en el ejemplo del empleado y que podemos llamar, siguiendo a Kane, “control plural”. Una decisión o una acción libre consistiría en el ejercicio de esa capacidad.

Sin embargo, la identificación de la libertad con la capacidad de elegir entre alternativas que se hallen al alcance o bajo el control efectivo del agente no parece plenamente correcta. Cuando las elecciones de un agente son absolutamente arbitrarias, cuando no hay modo de saber a qué se deben, de ver en ellas ninguna explicación racional, tendemos a considerarle, no como un agente libre, sino como un enfermo o un lunático. Y lo que parece estar ausente en este caso no es el control efectivo sobre las alternativas, sino cierto grado de control racional o volitivo-racional sobre la elección entre las mismas. Para ser libres, una decisión o una acción han de estar bajo el control volitivo-racional del agente, han de ser el fruto de su voluntad racional. De ahí que no consideremos las acciones realizadas en un estado de fuerte tensión emocional, que impide a un sujeto valorar las razones a favor y en contra de distintos cursos de acción y decidir de acuerdo con ellas, como ejemplos, o al menos como ejemplos claros, de acciones libres.

Sin embargo, el control volitivo-racional no resulta tampoco suficiente para que una decisión o una acción sean libres. Un sujeto puede ejercer ese control sobre sus decisiones y acciones sin ser un agente libre. Si, por ejemplo, los deseos, motivos y valores de un sujeto son el fruto de un proceso de lavado de cerebro o de un condicionamiento psicológico sistemático, o tal vez de una programación o incluso una manipulación directa de su cerebro, no consideramos a este sujeto como un agente libre, aun cuando conserve intactas sus capacidades de deliberación y elección racional. Y la razón de ello parece ser que no vemos a este sujeto como el genuino autor u origen de sus decisiones y acciones. El sujeto en cuestión, podríamos decir, no tiene el control último de las decisiones que toma y de las acciones que lleva a cabo como consecuencia de aquéllas. Así, además del control volitivo-racional sobre las decisiones y acciones, la libertad requiere también el control del agente sobre determinados factores, como valores, preferencias y creencias, sobre cuya base el agente toma esas decisiones y lleva a cabo

esas acciones. Aun cuando se trata sin duda de un requisito muy exigente, nuestras intuiciones sobre los casos mencionados de manipulación son muy claras al respecto y avalan esa clase de control, que podemos llamar control u origen último, como una condición genuina de la libertad y de su ejercicio.

Si consideramos la intencionalidad como una presuposición o un supuesto de base para la libertad, podemos decir, en un primer análisis, que el control plural o posesión de alternativas efectivamente accesibles al agente, el control volitivo-racional de la elección entre tales alternativas y el control último o control sobre los factores que dan cuenta de dicha elección son, plausiblemente, condiciones necesarias (y tal vez, en unión con ciertos requisitos de carácter cognitivo, también conjuntamente suficientes) de la libertad y de su ejercicio mediante decisiones y acciones libres.

3. *La responsabilidad moral*

La responsabilidad moral, por su parte, puede entenderse en varios sentidos. Algunos de ellos no han formado parte del debate tradicional sobre las relaciones de la responsabilidad moral con la libertad y el determinismo³. El sentido de la responsabilidad moral que ha estado en juego en este debate, y cuya compatibilidad con el determinismo ha sido objeto de controversia, puede expresarse como sigue: un agente es moralmente responsable de una acción suya si la ha llevado a cabo de tal manera que merece realmente reprobación si la acción era moralmente errónea y merece realmente crédito o tal vez alabanza si era moralmente correcta (cf. Pereboom 2007, p. 86). En una atribución sincera de responsabilidad moral a una persona por alguna acción gravemente perjudicial para otras personas, no dudamos de que esa persona merece realmente nuestra reprobación. Si nos indignamos sinceramente con una persona, estamos convencidos de que merece esa actitud nuestra hacia ella. Ahora bien –y éste es también un supuesto del debate tradicional– para que un agente sea moralmente responsable de una acción suya, para que merezca realmente alabanza o recriminación por ella, se requiere que la haya llevado a cabo con cierto grado de libertad o control, aun cuando hay discrepancia acerca de si todas las condiciones de la libertad que hemos distinguido (control plural, control

³ Uno de estos sentidos es el siguiente, en palabras de Pereboom: “Un agente puede considerarse moralmente responsable si es legítimo esperar de él que responda a preguntas tales como: ‘¿Por qué decidiste hacer eso? ¿Crees que hacer eso era lo correcto?’ y que evalúe críticamente lo que esas acciones indican sobre su carácter moral ... Los incompatibilistas no pensarían que ser responsable en este sentido de ‘responder’ es ni siquiera *prima facie* incompatible con el determinismo” (Pereboom 2007, p. 86). Scanlon (1988, 1998) mantiene una concepción de la responsabilidad moral muy cercana a ésta.

volitivo-racional, control último) son necesarias para la responsabilidad moral. Aunque, inicialmente, no parece justificado atribuir a un agente mérito o demérito por una acción que llevó a cabo sin que se cumpliera alguna de estas condiciones (como lo sugieren los ejemplos arriba aducidos), una reflexión más profunda puede arrojar dudas sobre la necesidad de alguna de ellas para la verdad o justificación de dicha atribución⁴.

4. *Compatibilismo e incompatibilismo*

El problema de la compatibilidad o incompatibilidad entre responsabilidad moral y determinismo surge sobre todo en relación con las condiciones anteriores. Si cada una de ellas es realmente necesaria para la responsabilidad moral, su incompatibilidad con el determinismo conllevará la incompatibilidad de éste con la responsabilidad moral. Así, la cuestión de la compatibilidad o incompatibilidad entre la responsabilidad moral y el determinismo puede concebirse como la cuestión de la necesidad de cada una de esas condiciones o aspectos de la libertad para la responsabilidad moral, así como de su compatibilidad o incompatibilidad con el determinismo.

Las posiciones tradicionales sobre esta cuestión siguen estando vivas en la actualidad. Podemos distinguir, en primer lugar, el incompatibilismo, la tesis según la cual el determinismo es incompatible con alguno(s) de los aspectos de la libertad necesario(s) para la responsabilidad moral y, por tanto, con esta última. Desde el punto de vista histórico, es la tesis más antigua. Y, por otra parte, parece también la posición más natural entre las personas no familiarizadas con la discusión filosófica. Fuera del ámbito filosófico, suele asumirse como algo obvio que la pregunta: “¿Somos libres o estamos determinados?”, plantea una disyunción exclusiva, de modo que la verdad de cada miembro de la disyunción implica la falsedad del otro. De este modo, el incompatibilismo adopta dos formas principales, dependiendo de cuál de los dos miembros se considera verdadero: el libertarismo, que defiende la existencia de la libertad y niega el determinismo, y el determinismo estricto, que adopta la actitud opuesta. Sin embargo, es también posible lo que cabría llamar incompatibilismo puro, que defiende simplemente la incompatibilidad entre ambos términos y se declara agnóstico acerca de cuál de los dos es verdadero.

⁴ Además de estas condiciones vinculadas a la libertad, la responsabilidad moral, sin embargo, requiere también condiciones adicionales de carácter cognitivo. Aquí, sin embargo, nos limitaremos a tratar de las primeras.

El incompatibilista puede basar su posición en la incompatibilidad del determinismo con la condición de alternativas o con la condición de origen o control último. Así, en la actualidad se distingue un incompatibilismo de alternativas o de margen (*leeway incompatibilism*) y un incompatibilismo de origen (*source incompatibilism*). También cabe (y éste es mi caso) ocupar ambas posiciones a la vez. Aunque la distinción entre ambas formas de incompatibilismo es reciente, es plausible considerar el determinismo estricto o fuerte, defendido por pensadores como Spinoza y Holbach, como una combinación, más o menos implícita, de ambos. No es habitual basar el incompatibilismo en la condición de control volitivo-racional, pues ésta suele considerarse –tal vez de forma un tanto precipitada– compatible con el determinismo. Y, de hecho, deterministas estrictos como Spinoza consideran la razón como la única forma de libertad de la que pueden gozar los seres humanos.

Es importante hacer notar que se puede aceptar a la vez la incompatibilidad del determinismo con las alternativas, e incluso con el control u origen último, y su compatibilidad con la responsabilidad moral, si se sostiene que tener alternativas, e incluso tener el control último de la acción, no es necesario para la responsabilidad moral. De hecho, esta posición es frecuente en la actualidad, especialmente a partir de la crítica de Harry Frankfurt (1969), que trataremos más abajo, a la necesidad de las posibilidades alternativas para la responsabilidad moral. John M. Fischer ha acuñado el término “semicompatibilismo” para designar esta posición, según la cual el determinismo es incompatible con la libertad (entendida en términos de alternativas o de control último), pero no con la responsabilidad moral. Ésta requiere libertad entendida como control volitivo-racional sobre la acción, pero esta condición es compatible con el determinismo.

Otra posibilidad teórica, ocupada de hecho en la actualidad por pensadores como Derk Pereboom, consiste en sostener, con Frankfurt o Fischer, que tener alternativas, aun siendo incompatible con el determinismo, no es necesario para la responsabilidad moral, pero que, a diferencia de Frankfurt o Fischer, tener el control último sobre la propia acción sí lo es y resulta también incompatible con el determinismo. Esta posición sostiene la incompatibilidad del determinismo con la libertad (como origen último) y con la responsabilidad moral. No es, pues, un semicompatibilismo sino un incompatibilismo completo. Además, según Pereboom, el control u origen último es también incompatible con el indeterminismo. De este modo, Pereboom caracteriza su propia posición como incompatibilismo estricto o fuerte, que constituye en realidad una forma de escepticismo

global sobre la libertad que sería necesaria para la responsabilidad moral y, por tanto, sobre esta última.

El libertarismo es, como hemos señalado, otra variedad importante del incompatibilismo. El libertarista sostiene, típicamente, que (al menos algunas de) las condiciones vinculadas a la libertad son necesarias para la responsabilidad moral e incompatibles con el determinismo. En su opinión, sin embargo, esas condiciones pueden satisfacerse si el determinismo es falso. Si, dado el pasado y las leyes de la naturaleza, el agente tiene efectivamente a su alcance al menos dos posibilidades alternativas, cabe considerarle como origen último de aquella que elige llevar a cabo; y, sea cual fuere la que elige, estaba en su poder haber optado por otra distinta. Así, desde este punto de vista, estaría justificado atribuirle responsabilidad, entendida como genuino mérito o demérito por la decisión que toma y la acción que lleva a cabo. Sin embargo, el libertarista tiene dificultades para dar cuenta de la condición de control volitivo-racional sobre la decisión y acción. Si, dado el conjunto de razones, rasgos de carácter, preferencias, valores y, en general, de todos los factores relevantes para tomar razonablemente una decisión, ésta continúa estando indeterminada, ¿cómo evitar la sospecha de que la decisión que de hecho acabe tomando el agente, sea cual fuere, será arbitraria, azarosa, mera cuestión de suerte y carente de justificación racional apropiada?⁵ ¿Y, en esas circunstancias, cómo podría estar realmente justificado considerarle como genuino merecedor de recriminación o alabanza, de castigo o premio? Bajo diversas denominaciones⁶, éste ha sido, y sigue siendo, uno de los principales problemas con los que se enfrenta el libertarismo.

En cuanto al compatibilismo, el espíritu general de esta posición puede expresarse diciendo que, para ella, la libertad no se opone a la determinación causal como tal, sino sólo a ciertas formas de la misma. Una acción libre no se distingue de una que no lo es por el hecho de que la primera, a diferencia de la segunda, no esté causalmente determinada, sino por el tipo de causas que determinan una y otra. Si la historia causal de la acción comprende factores tales como deseos, creencias, valores, preferencias y decisiones de un agente, la acción puede ser libre; si, en cambio, incluye factores como la coacción externa, la compulsión patológica o la manipulación neurológica, la acción

⁵ Una versión del libertarismo trata de llenar el hueco entre los factores previos a la decisión y esta misma apelando al *agente* mismo como causa. Sin embargo, esta forma de libertarismo recurre a un tipo de relación causal entre sustancias y eventos que, a diferencia de la relación estándar entre eventos, muchos consideran escasamente inteligible. Por otro lado, en mi opinión, la falta de control *racional* de la decisión sigue siendo un problema irresuelto para esta concepción.

⁶ Entre otras: “el problema del azar”, “la objeción de la arbitrariedad”, y también “la objeción de *Mind*”, por ser ésta la revista filosófica que, en un momento dado, publicó esta clase de críticas al libertarismo.

no será libre. El compatibilismo tuvo sus primeros proponentes en Thomas Hobbes y David Hume. En la actualidad es tal vez la posición que cuenta con un mayor número de adeptos, aun cuando una parte significativa de ellos son, en realidad, sólo semicompatibilistas. El compatibilismo tiene en la condición de control volitivo-racional su principal punto de apoyo, mientras que encuentra dificultades para dar cuenta de las condiciones de alternativas y de control u origen último. Y, ciertamente, si el determinismo es verdadero, no parece fácil explicar cómo puede haber posibilidades alternativas de decisión o acción efectivamente accesibles al agente, o cómo éste puede ser el origen último de su acción. Tradicionalmente, el compatibilismo ha afrontado esta dificultad tratando de mostrar, frente a las apariencias en contrario, que dichas condiciones, adecuadamente interpretadas, no son realmente incompatibles con el determinismo. Más recientemente, sobre todo a partir de las aportaciones de Harry Frankfurt y de semicompatibilistas como Fischer, la respuesta más frecuente del compatibilismo consiste más bien en negar que esas condiciones sean realmente necesarias para la responsabilidad moral. En esta versión, el compatibilismo tiene ante sí la tarea de ofrecer una concepción plausible de la libertad que sea reconocible como tal y que no incluya las condiciones indicadas. En su versión tradicional, la tarea consiste más bien en ofrecer un análisis de dichas condiciones que muestre su compatibilidad con el determinismo. Debido a esta compatibilidad, el análisis tenderá a adoptar un sesgo más o menos deflacionista.

Mientras que el incompatibilismo parece plantear condiciones muy exigentes para la responsabilidad moral, el problema del compatibilismo podría ser el contrario, a saber, su excesiva tolerancia a ese respecto. En otros términos, las condiciones propuestas por el incompatibilismo suscitan sospechas sobre su necesidad, mientras que las propuestas por el compatibilismo plantean más bien dudas sobre su suficiencia. Así, pues, las principales críticas por parte de los incompatibilistas insisten en la insuficiencia de los requisitos para la responsabilidad moral que propone el compatibilismo. Sin embargo, en correspondencia con su menor grado de exigencia, las condiciones compatibilistas parecen más fácilmente alcanzables para seres humanos como nosotros, mientras que, en especial desde la perspectiva compatibilista, las condiciones planteadas por el incompatibilismo parecen hallarse más allá de nuestro alcance, con lo que acaban favoreciendo el escepticismo sobre la libertad y la responsabilidad moral.

En las secciones que siguen analizaré, sucesivamente, las relaciones entre el determinismo y las posibilidades alternativas (sección II), entre las posibilidades

alternativas y la responsabilidad moral (sección III) y entre el determinismo, el control último y el control volitivo-racional (sección IV); finalmente, abordaré las dificultades asociadas al control último, con vistas a defender una versión propia del libertarismo (sección V).

II. DETERMINISMO Y POSIBILIDADES ALTERNATIVAS

En una primera aproximación, parece que la verdad del determinismo es obviamente incompatible con que haya alternativas de acción al alcance de una persona, ya que el determinismo es, en palabras de Ted Warfield, “la tesis según la cual la conjunción del pasado y las leyes implica todas las verdades” (Warfield 2000, p. 173). Esta definición es *grosso modo* equivalente a la de Van Inwagen que citamos más arriba, así como a esta otra, más breve, del propio Van Inwagen: “El *determinismo* ... es la tesis según la cual hay, en cualquier instante dado, exactamente un futuro físicamente posible” (1983, p. 3). Así, si el determinismo es verdadero, aunque varios desarrollos futuros del mundo son lógicamente posibles dado el pasado, sólo uno de ellos es físicamente posible, es decir, posible dadas también las leyes de la naturaleza. De este modo, parece que hemos de estar de acuerdo con Gary Watson cuando escribe: “Si el determinismo es verdadero, entonces claramente, en algún sentido, no hay posibilidades alternativas. Dadas las leyes de la naturaleza y las condiciones antecedentes, no es posible que alguien haga otra cosa que lo que de hecho hace” (Watson 1987, p. 154).

Las cosas, sin embargo, no son tan simples. Una expresión importante en el texto anterior es “en algún sentido”. ¿Qué se quiere decir exactamente cuando se afirma, en el contexto de las atribuciones de libertad y responsabilidad moral, que alguien habría podido hacer algo distinto de lo que hizo? Claramente no se pretende sostener sólo que era lógicamente posible una acción diferente; ni siquiera que una acción diferente era físicamente posible; prescindiendo del determinismo, ambas cosas pueden ser verdad, pero no resultan relevantes para justificar la atribución de responsabilidad moral. Lo que queremos decir es más bien que estaba efectivamente en poder o bajo el control del agente haber actuado de otro modo.

¿Nos priva el determinismo de ese poder o capacidad de hacer algo diferente de lo que hacemos? Vamos a suponer, como hemos sugerido más arriba, que las leyes de la naturaleza son deterministas, no probabilísticas, y que las propiedades mentales sobrevienen fuertemente a las propiedades físicas. Bajo estos supuestos, si el determinismo es verdadero y si mantenemos fijo el pasado hasta el instante mismo de la

acción, incluyendo la decisión del agente, parece que la respuesta a la pregunta anterior ha de ser afirmativa, a menos que el agente tenga poder para alterar el pasado o para violar alguna ley de la naturaleza. Pero parece claro que ninguna de las dos cosas está en poder de agentes finitos como nosotros⁷.

El razonamiento anterior guarda algunas semejanzas con un importante argumento destinado a mostrar la incompatibilidad del determinismo con la libertad entendida como el poder de actuar de manera diferente de como uno actúa (o control plural). Este argumento se conoce con el nombre de “Argumento de la Consecuencia”. En una presentación informal, debida a Van Inwagen, su principal defensor, el argumento es el siguiente:

Si el determinismo es verdadero, nuestros actos son las consecuencias de las leyes de la naturaleza y de sucesos en el pasado remoto. Pero no depende de nosotros lo que sucedió antes de que nacióramos, ni tampoco cuáles son las leyes de la naturaleza. Por lo tanto, las consecuencias de estas cosas (incluyendo nuestros actos presentes) no dependen de nosotros. (Van Inwagen 1983, p. 56)

El supuesto de la superveniencia fuerte o robusta de las propiedades mentales sobre las físicas es importante para la fuerza del argumento, pues sin él cabría la posibilidad de que tuviéramos poder sobre nuestras decisiones aunque no lo tuviéramos sobre los procesos físicos, incluyendo los movimientos físicos que en cada momento constituyen nuestras acciones⁸.

Un supuesto importante del argumento, que ha sido objeto de controversia, es el llamado “Principio de Transferencia de la Impotencia”, según el cual, si alguien no tiene elección (poder) sobre la verdad de cierta proposición ni sobre el hecho de que esa proposición implica lógicamente otra, no tiene tampoco elección o poder sobre la verdad de esta última. Este principio permite pasar de mi incapacidad de cambiar el pasado y las leyes naturales a mi incapacidad de hacer algo distinto de lo que en cada momento hago a través de mi incapacidad de alterar el hecho de que, dado el determinismo, el pasado y

⁷ La cuestión, sin embargo, está menos clara en relación con las leyes de la naturaleza. David Lewis sostuvo que hay algún sentido en el que cabe decir que podemos violar una ley natural (Lewis 1981). Para un intento reciente de cuestionar la inviolabilidad de (algunas) leyes naturales, cf. Berofsky 2009.

⁸ Sin embargo, aun sin ese supuesto, el argumento tendría ya consecuencias suficientemente graves, ya que, aceptando el determinismo físico, dejaría abierta la posibilidad de una desconexión sistemática entre nuestras decisiones y nuestras acciones. Obviamente, no queremos que nuestra libertad se limite a nuestras decisiones, sino que pretendemos que llegue también a nuestras acciones.

las leyes naturales implican todas las proposiciones verdaderas, incluidas las que describen las acciones que llevo a cabo. El principio es plausible, pero no incontrovertible. Sin embargo, aun cuando se mostrase falso, parece posible construir versiones del argumento que no descansan sobre él (cf. por ejemplo Fischer y Ravizza 1998, p. 22; Unger 2002, p. 5; Moya 2006, p. 27)

Parece claro que, junto con el supuesto de la superveniencia robusta, el argumento excluye las posibilidades alternativas entendidas en un sentido categórico, es decir, como posibilidades que un agente tiene de hacer B en *exactamente las mismas circunstancias* en que hizo A. Sin embargo, de acuerdo con una propuesta de George H. Moore⁹, lo que realmente queremos decir al afirmar que alguien habría podido hacer algo distinto de lo que hizo es que, si hubiera decidido (o elegido, o intentado) hacer algo distinto, lo habría hecho. Es ésta una interpretación condicional de las posibilidades alternativas de acción. La acción alternativa se hace aquí depender de un cambio en las circunstancias en las que un agente hizo A, a saber, de una decisión o elección distinta. Pero la verdad de la oración: “Si S hubiera decidido hacer algo distinto, lo habría hecho”, es compatible con la verdad del determinismo. Y si eso es lo que significa la oración “S habría podido hacer algo distinto”, la verdad de esta última es también compatible con el determinismo. Además, se puede mostrar, aunque no lo haré aquí, que, bajo esta interpretación condicional de las posibilidades alternativas, al menos una premisa del Argumento de la Consecuencia, que constituye una aplicación del Principio de Transferencia de la Impotencia, sería falsa¹⁰.

Sin embargo, el defensor de la incompatibilidad entre el determinismo y las posibilidades alternativas de acción está muy lejos de hallarse inerte frente a esta respuesta. En primer lugar, ya Roderick Chisholm (1964) arguyó convincentemente que “S habría podido hacer algo distinto” no es lógicamente (ni, *a fortiori*, semánticamente) equivalente a: “Si S hubiera decidido (elegido, intentado) hacer algo distinto, lo habría hecho”. En efecto, esta última oración podría ser verdadera y la primera falsa en el caso de que S no hubiera podido decidir o elegir algo distinto. Y hay ciertamente casos en los que, independientemente de la verdad del determinismo, una persona es incapaz de hacer cierta elección. Por ejemplo (cf. Fischer 2007, p. 50), debido a un trauma infantil con las serpientes, una persona podría ser incapaz de decidir seriamente tocar una serpiente. Es verdad de esta persona que, si decidiera tocar una serpiente, lo haría, pero es falso que

⁹ Según el testimonio de J. L. Austin (cf. Austin 1970)

¹⁰ Mostrar esto en detalle sería demasiado prolijo para un trabajo como éste. El lector interesado puede ver los detalles en Kane 1996, p. 47 o en Moya 2006, pp. 16-17

podría (voluntariamente) tocar una serpiente. El análisis condicional ha de presuponer, pues, para ser plausible, que S habría podido decidir algo distinto, pero “S habría podido decidir algo distinto” tiene la misma forma lógica que “S habría podido hacer algo distinto”, que es la oración que se pretendía analizar en términos condicionales, con lo que la propuesta no ha supuesto ningún avance, e incluso parece llevar implícito un regreso de actos *ad infinitum*.

El partidario de un análisis condicional puede contrarrestar estas objeciones apelando a un cambio en los factores que, desde la perspectiva del compatibilismo, dan lugar a decisiones libres, en particular las razones, entendidas como pares de creencias y deseos. Donald Davidson (cf. 1982, p. 147) adoptó esta estrategia. Según ella, “S habría podido decidir algo distinto” equivale a: “Si S hubiera tenido razones (deseos y creencias) distintas, habría decidido algo distinto”. En paralelo con la propuesta de Moore, la de Davidson ha de asumir también, para ser plausible, la verdad de “S habría podido tener razones (deseos y creencias) distintas”. Sin embargo, la propuesta no está ya expuesta a la objeción de que no supone ningún avance, ya que, a diferencia de “decidir”, “elegir” o “intentar”, “creer” o “desear” no designan acciones o actos; por la misma razón, la propuesta puede también evitar la amenaza de un regreso de actos *ad infinitum*. La contrapartida a estas ventajas, sin embargo, es, en mi opinión, una falacia de equivocidad en el uso del verbo “poder”. Así, en “S habría podido decidir algo distinto”, “habría podido” significa, *grosso modo*, “estaba en poder (o bajo el control) de S”, “dependía de S”; pero en “S habría podido tener razones (deseos y creencias) distintas”, “habría podido” significa más bien (aceptando el determinismo) “era lógicamente posible”. Pero la libertad de hacer o decidir otra cosa no equivale a la mera posibilidad lógica de hacer o decidir otra cosa, sino que alude más bien, como más arriba hemos señalado, a un poder o capacidad efectiva de obrar o decidir de otro modo.

A la vista de todo esto, las posibilidades alternativas de decisión o acción, como presunto requisito de la responsabilidad moral, deberían entenderse en sentido categórico. Al afirmar de alguien, en el contexto de una atribución de libertad y responsabilidad moral, que habría podido decidir algo distinto, queremos decir que, siendo como era y teniendo las creencias y deseos que tenía, estaba en su poder haber tomado una decisión distinta y, en consecuencia, haber obrado de otro modo. Y, aun cuando la cuestión dista mucho de estar definitivamente resuelta, la balanza se inclina hacia la incompatibilidad entre las posibilidades alternativas, entendidas categóricamente, y la verdad del determinismo.

III. POSIBILIDADES ALTERNATIVAS Y RESPONSABILIDAD MORAL

1. *El Principio de Posibilidades Alternativas (PPA) y los casos Frankfurt*

Como más arriba he indicado al exponer las distintas posiciones acerca del problema que nos ocupa, que el determinismo sea incompatible con las posibilidades alternativas, con la libertad o poder de obrar de otro modo, no implica sin más que sea incompatible con la responsabilidad moral. Se requiere también que las posibilidades alternativas sean necesarias para la responsabilidad moral. Es típico de muchos compatibilistas, y en especial de los semicompatibilistas, basar su posición en el rechazo de la necesidad de posibilidades alternativas para la responsabilidad moral, más que en la defensa de la compatibilidad entre el determinismo y las posibilidades alternativas.

La tesis según la cual la libertad, entendida como la capacidad efectiva de obrar de otro modo (o control plural), es necesaria para la responsabilidad moral, encuentra una expresión característica en el llamado “Principio de Posibilidades Alternativas” (PPA), que puede ser formulado como sigue: “Una persona es moralmente responsable de lo que ha hecho sólo si hubiera podido hacer otra cosa” (cf. Frankfurt 1969). Desde el punto de vista filosófico, la cuestión realmente importante no es el PPA como tal, sino la intuición central que subyace al mismo, a saber, que tener posibilidades alternativas de cierto tipo es necesario para la responsabilidad moral por lo que uno hace. En lo sucesivo, con la expresión “PPA” me referiré a esta intuición central. Por otra parte, limitaré la discusión sobre el PPA a la responsabilidad moral negativa, es decir, al merecimiento genuino de culpa o reprobación, dejando de lado la alabanza o aprobación¹¹.

Muchos filósofos (John Fischer, Derk Pereboom, Alfred Mele...) consideran hoy falso el PPA. Sus razones principales para ello tienen que ver con la existencia de supuestos contraejemplos al PPA llamados “casos Frankfurt”, debido a un caso presentado por Harry Frankfurt (1969). Los casos Frankfurt son ejemplos de situaciones conceptualmente posibles en las cuales, supuestamente, un agente es moralmente responsable de algo que ha hecho aun cuando, debido a circunstancias de las que no es consciente, carecía de alternativas del tipo apropiado. Por mi parte, pienso que estos casos no son verdaderos contraejemplos al PPA, un principio que tiendo a considerar verdadero, al menos en relación con la culpabilidad moral.

¹¹ En trabajos recientes (Moya 2010 y Moya, en prensa) he defendido una posición asimétrica, según la cual se requieren alternativas para merecer culpa o reprobación, pero no para merecer alabanza o aprobación por una acción.

Un caso Frankfurt típico sería el siguiente:

(MENTIRA): Inés, una estudiante, odia a su compañero Antonio y decide mentirle sobre la hora de un examen inminente. Como resultado de ello, Antonio no realiza el examen, con graves consecuencias para él. Sin que Inés tenga la menor sospecha, Black, un neurocirujano malintencionado, que también quiere que Inés mienta a Antonio, ha implantado en el cerebro de ésta un dispositivo que le permite seguir su deliberación. Si, sobre esta base, resulta claro para Black que Inés decidirá no mentir, pulsará entonces un botón que asegurará que Inés acabará decidiendo mentir. Sin embargo, Black prefiere no intervenir si no es indispensable. Y no lo es, porque Inés, deliberando por sí misma y teniendo en cuenta sus propias razones, decide mentir, y miente, a Antonio sin que Black tenga que pulsar el botón ni intervenir en modo alguno.

El ejemplo suscita fuertemente la intuición de que Inés es moralmente culpable de su decisión y acción de mentir a Antonio, aun cuando, debido a la presencia oculta de Black, no hubiera podido siquiera decidir no mentirle. Se trata, pues, aparentemente, de un contraejemplo al PPA.

Frente a ejemplos de este tipo, cabría responder, en defensa del PPA, que Inés no carecía de alternativas. Hubiera podido mostrar la señal que habría alertado a Black de que iba a decidir no mentir a Antonio. Asimismo, y como consecuencia de ello, aunque habría acabado mintiendo debido a la intervención de Black, no habría mentido “por sí misma” (Naylor 1984) o intencionalmente (Davidson 1982). Sin embargo, los adversarios del PPA pueden construir casos en los que la señal de alerta podría ser un mero suceso, más allá del control del agente, como un sonrojo o algún sutil cambio neurológico (cf. Fischer 1994). Y, en este contexto, no mentir por sí misma o no mentir intencionalmente serían meras consecuencias de esos sucesos, imprevisibles para Inés y ajenas a su control. Pero no resulta plausible que apuntar a alternativas de este tipo baste para salvaguardar el PPA. El espíritu del PPA se manifiesta en juicios como éste: el agente fue culpable de lo que hizo porque, estando en su poder haber actuado de un modo moralmente más correcto, no lo hizo. Las alternativas que, en el marco de este principio, se consideran *necesarias* para que el agente sea moralmente culpable de lo que hizo han de ser moralmente significativas: formas de actuar (incluyendo la omisión de lo que se hizo) que se hallaban bajo el control del agente y que, de haber optado por ellas, habrían modificado

sustancialmente su responsabilidad moral por lo que hizo. Alternativas como sonrojarse o sufrir un cambio neurológico, siendo meros sucesos, no se ajustan a este criterio. Incluso en el terreno de las acciones, Inés hubiera podido mentir a Antonio en voz más baja o más alta, o usando palabras distintas, pero éstas no eran alternativas relevantes para su responsabilidad moral: de haberlas llevado a cabo, ella habría sido tan culpable de mentir de Antonio como lo es ahora. Y alternativas moralmente significativas o robustas (por usar el término introducido por Fischer), como decidir no mentir y cumplir esa decisión, no estaban al alcance de Inés.

Según Frankfurt, la falta de alternativas sólo es relevante para la responsabilidad moral cuando el sujeto actúa como lo hace *porque* no tiene alternativas, como sucedía en nuestro ejemplo del empleado de banco víctima de un atraco. Pero éste no es el caso de Inés. Ella habría actuado como lo hizo aun cuando Black no hubiera estado al acecho y ella hubiera podido actuar de otro modo. Inés fue culpable de mentir a Antonio porque lo decidió por sí misma, a través de su propia deliberación y por sus propias razones; no importa que no hubiera podido siquiera decidir no mentirle, ya que este hecho no tuvo papel alguno en su decisión y acción de mentir. Así, Frankfurt y sus seguidores sostienen que la responsabilidad moral de un agente está determinada exclusivamente por rasgos de la *secuencia real* que condujo a la decisión y la acción.

MENTIRA es, podríamos decir, un caso Frankfurt clásico. Pero este tipo de casos se enfrentan a una importante objeción, planteada por filósofos como Robert Kane (1985), Carl Ginet (1996) y, sobre todo, David Widerker (1995). La objeción, que suele llamarse “la objeción del dilema”, puede plantearse como sigue. Pensemos en la señal que Black emplea para *no* intervenir, como, por ejemplo, que Inés no se sonroje en cierto momento. Con respecto a esta señal, cabe plantear el siguiente dilema: o bien la señal es (o está asociada con) una condición causalmente suficiente para la decisión de Inés de mentir a Antonio, o bien no lo es (no lo está). En el primer caso, la decisión es inevitable, pues está causalmente determinada, pero entonces un incompatibilista no aceptará que Inés es moralmente responsable de ella. En el segundo caso, si la señal es meramente un signo fiable, pero no infalible, de lo que Inés decidirá, Inés puede ser moralmente responsable de su decisión y acción, pero no hay razón para aceptar que no hubiera podido decidir no mentir. En uno u otro cuerno del dilema, el PPA permanece indemne.

2. *Nuevas críticas al PPA. Y una nueva defensa.*

Como consecuencia de la objeción del dilema, los seguidores de Frankfurt se han visto obligados a modificar los ejemplos para evitarla. La estrategia más común consiste en optar por la segunda rama del dilema y asumir que la decisión no está causalmente determinada; ello deja abiertas algunas posibilidades alternativas; sin embargo –y éste es un aspecto crucial– se trata de lograr que ninguna de esas posibilidades sea robusta o moralmente significativa, con lo que su presencia no bastará para salvaguardar el PPA. Presento a continuación un caso Frankfurt de mi invención con estas características. El ejemplo combina rasgos de algunos casos propuestos recientemente¹²:

(MEZQUINO): Roberto es un empresario de éxito, con una situación financiera desahogada. Su hermano Rodrigo, en cambio, está pasando apuros económicos. Un día, Rodrigo visita a Roberto para decirle que tiene algunas deudas, no muy elevadas, pero que debe pagar en unos pocos días si no quiere tener serios problemas. Así, le pide a Roberto que le preste algún dinero para cancelarlas. Roberto es una persona tacaña y egoísta, aunque no hasta el punto de no poder actuar también de manera altruista si las razones para ello son importantes. Dada su forma de ser, en estas circunstancias Roberto sólo podría decidir prestar a Rodrigo el dinero que le pide si tomara en serio razones morales para hacerlo, cosa que está en su mano hacer. Pero tomar en serio estas razones no es suficiente para que decida ayudar a Rodrigo; aunque tomara en serio estas razones, Roberto podría negarse a ayudar a su hermano por razones de interés propio, vinculadas a su carácter mezquino y egoísta. No obstante, sin él saberlo, Roberto no podría tomar en ese momento la decisión de ayudar a Rodrigo, pues su aparato circulatorio se halla momentáneamente en un estado tal que, si tomase en serio razones morales para ayudar a Rodrigo, y con ellas la posibilidad de hacerlo, sufriría un ataque cardíaco que provocaría en él un estado de coma, lo que le impediría prestar a su hermano el dinero que tanto necesita. Sin embargo, Roberto no sufre ataque cardíaco alguno, ya que no considera seriamente razones morales para echar una mano a su hermano y se niega a prestarle el dinero, con la excusa de que él mismo está atravesando momentos económicamente difíciles.

¹² Como el ejemplo de Derk Pereboom *Evasión de impuestos* (Pereboom 2009) y el de David Widerker *Disfunción-cerebral-W* (Widerker 2006),

Advirtamos que, en este ejemplo, la decisión que de hecho toma Roberto no está causalmente determinada, aunque, sin él saberlo, no habría podido tomar una decisión distinta. Pero, según los seguidores de Frankfurt, ello no obsta para que sea moralmente responsable de la decisión que tomó. Los defensores del PPA argüirán que Roberto tenía alternativas: hubiera podido tomar en serio razones morales para ayudar a Rodrigo. Sus oponentes admitirán que Roberto tenía esta alternativa, pero sostendrán que no era una alternativa robusta. Es cierto que, si la hubiera escogido, se habría librado de la culpa que ahora tiene por no ayudar a Rodrigo, pero ello habría ocurrido por pura suerte, de un modo ajeno a su control, ya que, ignorante como era de su condición física en ese momento, Roberto no podía creer razonablemente que sólo por tomar en serio razones morales se libraría de culpa. Recordemos que (prescindiendo de su pasajera condición física) tomar en serio esas razones era plenamente compatible con su decisión de no ayudar a Rodrigo.

Sin embargo, creo que cabe defender plausiblemente el PPA frente a este tipo de argumentos y frente a supuestos contraejemplos como MEZQUINO. Pasemos a ello.

Es cierto que, en general, tomar en serio razones morales en contra de una decisión o acción moralmente objetable no es una alternativa robusta. En circunstancias normales, sólo lo sería evitar esa decisión o acción. Sin embargo, con independencia de la discusión sobre los casos Frankfurt y el PPA, nuestros juicios sobre el carácter de una alternativa que un agente tiene y sus efectos sobre su responsabilidad moral son altamente dependientes del contexto, y en particular de las alternativas que, en dicho contexto, se hallan de hecho en poder de dicho agente¹³. Pensemos en el siguiente ejemplo:

(CARTERISTA): Al volver del trabajo, Evaristo se encuentra con Ernesto, un viejo amigo suyo. Ernesto le dice que necesita urgentemente tomar un taxi para llegar a tiempo a una importante reunión y le pide prestada una pequeña cantidad, ya que se ha dado cuenta de que no lleva consigo dinero suficiente para el taxi. Desde luego, continúa, le devolverá el dinero lo antes posible. Evaristo cree que puede fácilmente hacer ese favor a Ernesto, pero no está de buen humor y no tiene ganas de ser amable. Así, le responde que, sintiéndolo mucho, no puede ayudarle, ya que

¹³ La argumentación que sigue procede de algunos trabajos míos, en especial Moya 2010 y 2011. Su germen se halla ya, sin embargo, en Moya 2009. Un importante desarrollo y perfeccionamiento de esta defensa del PPA puede verse en Moya 2014; en este artículo ofrezco una respuesta a algunas objeciones a dicha defensa, en especial las planteadas por Pereboom en su artículo de 2012.

esa mañana olvidó coger su cartera al salir de casa y no lleva dinero encima. Sin él saberlo, sin embargo, de hecho, no lleva encima dinero ni cartera, pues un carterista se la robó media hora antes mientras viajaba en el metro.

Si, posteriormente, alguien le criticase por no prestar dinero a Ernesto, Evaristo podría excusarse diciendo que no era justo culparle por ello, ya que no hubiera podido prestarle ese dinero aunque hubiera querido hacerlo. Pero Evaristo no está libre de culpa por completo. Lo habría estado si hubiera intentado sinceramente ayudar a Ernesto llevándose la mano al bolsillo (vacío) de su americana. La razón por la que hacer tal cosa le habría librado de la culpa que ahora tiene es que eso era todo lo que razonablemente podía haber hecho en las circunstancias del caso para comportarse decentemente con su amigo. Su intención sincera de ayudarlo llevándose la mano al bolsillo, era, *en esas circunstancias*, una alternativa robusta y eximente.

Imaginemos ahora una variación en el ejemplo. Supongamos que el robo no ha tenido lugar y que Evaristo tiene la cartera en su bolsillo. En estas circunstancias, intentar ayudar a Ernesto llevándose la mano a la cartera ya no hubiera sido una alternativa robusta y eximente. Para librarse de culpa, Evaristo tendría que haber prestado a Ernesto el dinero que éste necesitaba, ya que ahora podía haberlo hecho. Un mismo tipo de acto puede ser una alternativa robusta en una situación y no en otra, dependiendo de lo que el agente pueda realmente hacer en cada una de ellas.

Nuestras reacciones ante estos ejemplos muestran que nuestros juicios sobre la responsabilidad moral de las personas no dependen solamente, como sostienen Frankfurt y sus seguidores, de la secuencia real que conduce a la acción, sino también de rasgos modales de cada situación, y en concreto de lo que un agente *puede* realmente llevar a cabo, de las alternativas que están de hecho a su alcance en una situación determinada.

Un modo de explicar nuestros juicios sobre CARTERISTA y su variación es verlos como la aplicación de un principio que, en nuestra opinión, forma parte del esquema conceptual del sentido común y que conecta los conceptos de poder, obligación y responsabilidad moral:

(POR) Si alguien no puede razonablemente hacer más de lo que hace para cumplir con su deber, no está obligado a hacer más, de modo que no es moralmente responsable de no hacerlo.

El principio POR es en realidad la conjunción de dos condicionales. El primero (“Si alguien no puede... no está obligado...”) es la contraposición de un viejo principio moral de raigambre kantiana: “deber” implica “poder”. El segundo (“si no está obligado... no es moralmente responsable...”) es una aplicación de la tesis, ciertamente plausible, según la cual la culpabilidad moral por hacer (no hacer) algo requiere la obligación moral de no hacerlo (hacerlo).

Armados con este bagaje conceptual, podemos ahora sostener plausiblemente, frente a los seguidores de Frankfurt, que, en MEZQUINO, tomar en serio razones morales para ayudar a su hermano era una alternativa robusta y exigente que Roberto tenía, ya que, en la situación en la que se hallaba, no hubiera podido hacer más para cumplir con su deber de ayudar a su hermano Rodrigo.

De este modo, casos Frankfurt tan complejos como MEZQUINO no prueban la falsedad del PPA. En este ejemplo, hay algo moralmente relevante que Roberto tenía la obligación de hacer, no hizo y pudo haber hecho, de modo que su responsabilidad moral no se produce en ausencia de alternativas robustas y se debe, en parte, a la presencia de tales alternativas. Roberto debió ser sensible a la situación y necesidades de su hermano y haberse puesto en su lugar, considerando seriamente la posibilidad de ayudarlo e intentando sinceramente hacerlo, aun cuando ese intento, sin culpa alguna por su parte, no habría tenido finalmente éxito.

3. Alternativas y capacidad de respuesta a razones

Alguien podría responder a esta defensa del PPA tratando de eliminar alternativas moralmente significativas, como la que Roberto tiene en MEZQUINO, haciendo que el acceso a estas alternativas dependa, sin que el sujeto lo sepa, de una condición moralmente irrelevante en sí misma y sobre la cual no tiene control. Supongamos, por ejemplo, que Roberto está constituido (sin que él lo sepa) de tal manera que, en la situación en que se halla, sólo podría tomar en serio razones morales para ayudar a su hermano si un pensamiento sobre (digamos) el vuelo de las palomas cruzara antes por su mente¹⁴. Alternativamente, cabría suponer, en la estela de Widerker (2006), que, tras la petición de Rodrigo, o tal vez un poco antes, deja de darse en el cerebro de Roberto una condición neurológica necesaria para que tome en serio razones morales. No sería correcto ahora decir que estaba en poder de Roberto tomar en serio razones morales, ya

¹⁴ De hecho, Neil Levy argumentó en este sentido en una contribución al debate sobre mi defensa del PPA en el *blog* “Flickers of Freedom”.

que ello dependía de algo sobre lo que no tenía control, a saber, un pensamiento casual sobre el vuelo de las palomas o un cambio neurológico. Sin embargo, Roberto es responsable de no ayudar a Rodrigo, siguiendo la propuesta de Frankfurt, porque rechazó ayudarlo por sus propios motivos egoístas, sin que nada ni nadie le obligara y sin tener una buena justificación para hacerlo, a pesar de carecer de cualquier alternativa robusta.

Un defensor del PPA puede no obstante responder a esta estrategia negándose a aceptar la responsabilidad moral del agente en tales condiciones, no, como sucede en la defensa del dilema, sobre la base de que la decisión está causalmente determinada, sino apelando al hecho de que, en esas circunstancias, la capacidad del agente de responder a razones, una capacidad ampliamente considerada, tanto por incompatibilistas como por compatibilistas, como una condición necesaria de la responsabilidad moral, se halla seriamente dañada¹⁵. *Grosso modo*, en términos de Fischer, al deliberar y decidir, un agente satisface esta condición en la medida en que, manteniendo constante el “mecanismo” de razonamiento práctico que emplea en la situación en cuestión, hay situaciones posibles en las que existen razones suficientes para decidir y actuar de otro modo y en las que el agente reconoce esas razones y decide y actúa de otro modo (cf. Fischer 1994, Fischer y Ravizza 1998). Según esto, si, con el mismo “mecanismo” de deliberación que ahora emplea, un agente seguiría tomando la misma decisión que ahora toma aun cuando se le presentaran razones decisivas para decidir de otro modo, el agente no satisface la condición en cuestión; sus capacidades para el razonamiento práctico están deterioradas, lo que arroja serias dudas sobre su responsabilidad moral por la decisión que ahora toma. Apliquemos ahora estas consideraciones a nuestro ejemplo.

En el caso de Roberto, reconstruido según la nueva estrategia, pensemos que, según está constituido psicológicamente, por muy buenas razones que pudiera contemplar para ayudar a Rodrigo, nunca las tomaría en serio ni respondería a ellas a menos que, casualmente, se le ocurriera pensar en el vuelo de las palomas. Pero asumimos, plausiblemente, que un agente moral normal debería poder tomar en serio y responder directamente a consideraciones morales, sin que tenga que mediar la ocurrencia casual de un pensamiento como ése. Así, en esta variación del ejemplo, se consigue eliminar alternativas robustas al precio de hacer de Roberto un agente con graves defectos en su “mecanismo” de razonamiento práctico y toma de decisiones, lo que arroja serias dudas sobre su responsabilidad moral. Algo semejante cabe decir del caso en que está ausente

¹⁵ He defendido este tipo de respuesta en Moya 2003, 2006, cap. 2 y 2007.

una condición neurológica causalmente necesaria para que tome en serio razones morales y, con ello, para que decida de modo moralmente correcto. En estas condiciones, el agente sería incapaz de decidir de modo moralmente correcto por muy fuertes que fuesen las razones que se le ofrecieran o en las que pudiera pensar para decidir de ese modo, ya que nunca las tomaría en serio. De nuevo, la capacidad de responder a razones se ve seriamente dañada, lo que afecta a las intuiciones sobre la responsabilidad moral del agente por su decisión. El PPA se mantiene, pues, incólume.

Cabría responder a la objeción anterior sosteniendo que las condiciones moralmente irrelevantes que son necesarias para que Roberto tome en serio razones morales no son realmente parte integrante de su “mecanismo” de razonamiento práctico, sino rasgos meramente accidentales que sólo se dan en la situación real y de los que es lícito prescindir para evaluar la capacidad de dicho “mecanismo” para responder a razones¹⁶. Así, cabría asumir que, con el “mecanismo” que ahora emplea, si las razones para ayudar a Rodrigo fuesen más fuertes de lo que son (por ejemplo, si Rodrigo estuviese en peligro de perder su casa si no devolvía su deuda), Roberto tomaría en serio esas razones morales más fuertes sin necesidad de pensar en el vuelo de las palomas. En el caso del cambio neurológico, cabe suponer que esas razones más poderosas activarían de nuevo la condición causalmente necesaria para que las tomara en serio.¹⁷ En la situación real, sin embargo, no está en poder de Roberto tomar en serio las razones morales más débiles de las que es consciente, pues el pensamiento sobre las palomas o la condición neurológica, ahora ausentes, son aquí factores causalmente necesarios para que las tome en serio.

La propuesta es interesante, ya que, al parecer, consigue eliminar alternativas robustas respetando la capacidad del “mecanismo” de deliberación y decisión del agente para responder a razones. Sin embargo, la dialéctica no termina aquí. El problema, según lo veo, es ahora el siguiente: si se asume que el “mecanismo” de deliberación y decisión que Roberto emplea en la situación real es capaz de responder a razones y de tomar en serio razones morales más poderosas, es entonces muy difícil aceptar que sea, literalmente, causalmente imposible para él, en la situación real, a menos que casualmente

¹⁶ Debo esta observación a Manuel Vargas, en conversación. Lo cierto es que Fischer no proporciona criterios claros de la identidad de un “mecanismo” de deliberación.

¹⁷ Derk Pereboom es autor de una propuesta de este tipo, en respuesta a distintas objeciones, incluida la mía. Cf. su ejemplo *Tax Cut*, en Pereboom 2013. Cf. también su ejemplo *Tax Evasion (3)*, en Pereboom 2009, p. 117. Una respuesta muy cercana a ésta me fue también planteada verbalmente por Tobies Grimaltos.

piense en las palomas, tomar en serio las razones morales más débiles que de hecho tiene en cuenta, con vistas a decidir según ellas. En cuanto a la otra variante del ejemplo, si el cambio neurológico es reversible en virtud de razones más fuertes,¹⁸ no vemos por qué no pueda serlo en virtud de razones más débiles, si Roberto se esfuerza en prestarles la debida atención. Una cosa es decir que, dado su carácter egoísta, es improbable que Roberto tome en serio esas razones morales y les dé preeminencia sobre razones de interés propio, y otra cosa muy distinta es decir que (en ausencia del pensamiento o la condición neurológica en cuestión) es causalmente imposible que lo haga. ¿Por qué tendríamos que aceptar esto último? Con frecuencia, atender seriamente y dar preeminencia a razones morales sobre motivos de interés propio, con vistas a orientar las decisiones según ellas, requiere un mayor esfuerzo de voluntad que hacerlo a la inversa, pero esto no significa que esté más allá del poder de un agente, si su “mecanismo” de razonamiento práctico es capaz de responder a razones morales. Roberto debería haber llevado a cabo ese esfuerzo y haber tomado en serio las razones morales de las que era consciente, con vistas a decidir de manera moralmente correcta. Aun en el caso de que, debido a su pasajera condición física, no hubiera podido llegar a decidir así, esa atención sincera a razones morales para hacerlo era, dado el principio POR, una alternativa robusta. Y puesto que, según mi oponente, su “mecanismo” de deliberación estaba intacto, no hay razones claras para aceptar que no estaba a su alcance.¹⁹ Estipular que no lo estaba no es suficiente para que el ejemplo sea convincente.

Por todo lo que llevamos dicho, hemos de dejar abierta la cuestión de la verdad del PPA. Aunque, por los argumentos anteriores, ninguno de los contraejemplos propuestos hasta ahora está, en mi opinión, libre de problemas, no podemos descartar que se pueda diseñar en el futuro alguno que lo esté. Al menos, he ofrecido razones para pensar que la tarea no será en absoluto sencilla. Y es sintomático que ésta sea la situación más de cuarenta años después de que Frankfurt ofreciera su primer (supuesto) contraejemplo al PPA.

¹⁸ Si no lo es, el ejemplo sucumbe a la objeción de que el mecanismo de deliberación y decisión no es capaz de responder a razones.

¹⁹ Excepto asumiendo el determinismo. Pero ello haría el ejemplo vulnerable a la objeción del dilema y constituiría una petición de principio contra el incompatibilista.

IV. DETERMINISMO Y CONTROL

1. *Control último, control volitivo-racional y compatibilismo*

Además de la condición de alternativas, la condición de control último sobre las propias decisiones y acciones ha sido también objeto de un encendido debate. Este debate se ha desarrollado en varias dimensiones. En primer lugar, se ha discutido sobre el contenido de dicha condición y, en relación con ello, sobre la posibilidad de satisfacerla. En segundo lugar, la discusión ha girado en torno a su necesidad para la responsabilidad moral. Finalmente, y en conexión con la cuestión del contenido de dicha condición, ha sido también objeto de controversia su compatibilidad con el determinismo.

Como hemos visto, las intuiciones sobre la importancia del control último para la responsabilidad moral derivan, en gran parte, de la convicción, ampliamente compartida, según la cual no está justificado considerar a un agente como autor u origen genuino de una acción en determinadas circunstancias, que incluyen la coacción externa, la compulsión patológica o la manipulación directa del cerebro. Sin este terreno común, la discusión no podría siquiera comenzar. Pero el terreno es exiguo y las divergencias comienzan en seguida. Así, por ejemplo, no hay acuerdo sobre si la programación o el condicionamiento psicológico, e incluso determinadas formas de manipulación neurológica, conllevan también la pérdida de la autoría u origen genuino. La discrepancia mayor, sin embargo, concierne a la cuestión de si el determinismo es equiparable, por lo que respecta a sus efectos sobre la autoría u origen, y con ello sobre la responsabilidad moral, a los factores que forman parte del terreno común. Característicamente, los incompatibilistas responden afirmativamente, mientras que la respuesta de los compatibilistas es negativa. Estas respuestas dependen también, sin embargo, del contenido que unos y otros dan a esta condición, de modo que no es claro que sean respuestas a la misma pregunta.

La noción de control último aúna las ideas de control y ultimidad de origen. Para ser origen o fuente genuina de una decisión o acción, de modo que esté justificado atribuirle responsabilidad moral por ella, el agente ha de haber tenido un control adecuado del proceso que le ha llevado a esa decisión o acción. Podemos quizá servirnos, para comenzar la discusión, de la caracterización que Derk Pereboom hace del requisito de origen (O), ya que es lo bastante general como para resultar aceptable para la mayor parte de los participantes en el debate, si no por todos. La formulación de Pereboom se restringe a la decisión, pero puede extenderse también a la acción:

Si un agente es moralmente responsable por decidir realizar una acción, entonces la producción de esta decisión debe ser algo sobre lo cual el agente tiene control; y el agente no es moralmente responsable por la decisión si es producto de una fuente sobre la cual no tiene control. (Pereboom 2001, p. 4)

El control sobre el proceso de decisión se requiere para que el agente sea su autor u origen de modo que esté justificado hacerle responsable de ella. Más allá de esto, sin embargo, hay serias discrepancias acerca de qué supone tener control sobre la decisión, no sólo entre compatibilistas e incompatibilistas, sino también en el interior de uno y otro bando.

En el campo del compatibilismo, podemos encontrar distintas formas de concebir el control que se requiere para la autoría genuina y la responsabilidad moral. El desarrollo histórico del compatibilismo puede entenderse como una serie de intentos sucesivos de entender ese control, cada uno de los cuales trata de remediar los defectos del anterior.

El compatibilismo anterior a Harry Frankfurt, que incluye autores como Th. Hobbes, D. Hume, J. S. Mill y A. J. Ayer y que podemos denominar, siguiendo a Watson, compatibilismo clásico, tiende a considerar una forma de control volitivo-racional de la acción (que incluye posibilidades alternativas concebidas en términos condicionales) como necesaria y suficiente para la responsabilidad moral. *Grosso modo*, según esta posición, la libertad es la capacidad de hacer lo que uno desea y decide hacer, la capacidad de actuar según las propias razones, según las propias preferencias, deseos y decisiones. Y actuar libremente consiste en ejercer esa capacidad, haciendo efectivamente lo que uno quiere y decide hacer. El agente controla su acción a través de sus propios deseos y decisiones. En esta perspectiva, las intuiciones relativas al origen de la acción se satisfacen en la medida en que las propias preferencias, motivos y razones son parte de nuestro carácter y personalidad, de nosotros mismos como entidades psicológicas, de modo que, al decidir y actuar por tales factores, somos nosotros mismos quienes damos origen a nuestras acciones. En esta perspectiva, el control sobre la acción se pierde cuando sus causas no son los factores indicados, sino más bien factores como la coacción externa o una compulsión irresistible de tipo patológico. Parece claro que la libertad así concebida es compatible con el determinismo. Lo decisivo para la libertad no es que nuestras acciones carezcan de causas, sino que estén causadas por nuestros deseos, motivos y preferencias, no importa si éstos a su vez, tienen causas que se extienden indefinidamente hacia el pasado

No es difícil darse cuenta de que la libertad o control así concebido resulta una condición demasiado débil y, así, insuficiente para fundamentar la responsabilidad moral si pensamos que, por ejemplo, niños pequeños, e incluso animales superiores, podrían satisfacerla.

Según Harry Frankfurt, cuya obra ha sido de crucial importancia en el debate contemporáneo sobre la responsabilidad moral, la libertad, tal como la concibe el compatibilismo clásico, corresponde a lo que cabe llamar libertad de acción, no a la libertad de la voluntad o libre albedrío. Pero es este último concepto el que ha constituido el foco principal del debate filosófico. Según Frankfurt, actuar por nuestros propios deseos o decisiones no es suficiente para que se nos pueda considerar justificadamente como autores genuinos y responsables de nuestras acciones. Se requiere además que nuestros deseos de actuar de cierta manera, que Frankfurt llama deseos de primer orden, estén también sometidos a nuestro control volitivo. Y tenemos control sobre esos deseos cuando son los deseos que deseamos tener. Desear tener o no tener cierto deseo de primer orden es un deseo de segundo orden. No todos los deseos de primer orden que tenemos nos llevan efectivamente a actuar; aquellos que sí lo hacen son, en la terminología de Frankfurt, nuestra voluntad. Y, entre nuestros deseos de segundo orden, aquellos cuyo objeto es nuestra voluntad son voliciones de segundo orden. No son meros deseos de tener o no tener ciertos deseos, sino deseos de que ciertos deseos nos muevan o no a actuar. Sobre esta base podemos decir, siguiendo a Frankfurt, que, del mismo modo que la libertad de acción es la capacidad de hacer lo que deseamos hacer, la libertad de la voluntad es la capacidad de tener la voluntad que deseamos tener (cf. Frankfurt 1971, pp. 20-21). Y, del mismo modo que ejercemos nuestra libertad de acción cuando hacemos lo que deseamos hacer, ejercemos nuestra libertad de la voluntad cuando tenemos la voluntad que deseamos tener, es decir, cuando actuamos por deseos que aprobamos y por los que deseamos ser movidos a actuar. En este caso, actuamos libremente, o por nuestra libre voluntad:

Supongamos que una persona ha hecho lo que quería hacer, que lo ha hecho porque quería hacerlo y que la voluntad por la que fue movido cuando lo hizo era su voluntad porque era la voluntad que quería tener. En ese caso, hizo tal cosa libremente y por su propia libre voluntad. (Frankfurt 1971, p. 24)

Actuar libremente y por nuestra libre voluntad es todo lo que se requiere para que seamos autores genuinos de nuestras acciones y tengamos sobre ellas el control apropiado para ser moralmente responsables de ellas. Nuestras voliciones de segundo orden son parte esencial de nuestro yo personal, de modo que cuando actuamos por deseos con los que nos identificamos a través de tales voliciones, somos nosotros mismos el origen de esas acciones.

Para la responsabilidad moral por nuestras decisiones o acciones no se requiere que pudiéramos haber decidido o actuado de otro modo²⁰. Como en el caso del compatibilismo clásico, el control que Frankfurt exige para la responsabilidad moral es compatible con el determinismo. Para ser moralmente responsables de nuestras acciones es necesario tener voliciones de segundo orden, pero no es necesario que éstas estén causalmente indeterminadas. Además, para Frankfurt, cuál sea la historia u origen causal de nuestras voliciones de segundo orden, que forman el núcleo de nuestro yo, es irrelevante para la responsabilidad moral por nuestras acciones.

Para otros autores compatibilistas, como Gary Watson o Susan Wolf, el núcleo central del yo está constituido por el sistema de valores, más que por las voliciones de segundo orden. Se trata de concepciones del yo más intelectualistas y menos voluntaristas que la de Frankfurt. Watson distingue entre dos fuentes distintas e independientes de motivación: los deseos y los valores. Podemos desear cosas que no consideramos valiosas y, a la inversa, podemos valorar cosas que no logramos desear, en el sentido de vernos motivados realmente a alcanzarlas. Esta distinción entre deseos (pasiones, en terminología humeana) y valores se halla en la raíz del problema del libre albedrío. El libre albedrío puede concebirse, en la perspectiva de Watson, como la capacidad de desear y actuar de acuerdo con nuestro sistema de valores. Y ejercemos esa capacidad al lograr la armonía entre nuestro sistema de valores y nuestro sistema motivacional. No conseguimos ejercerlo, en cambio, si lo que nos motiva realmente a actuar son deseos a favor de cosas o acciones que no consideramos valiosas. Al actuar por nuestros valores, que constituyen el núcleo central de nuestro yo, damos origen a nuestras acciones y somos, pues, responsables de ellas. Así, como en las propuestas anteriores, se intenta satisfacer las intuiciones asociadas al requisito de origen o control último a través del control volitivo-racional de un agente sobre sus decisiones y acciones. Mientras que, en el caso del compatibilismo clásico, este control consiste en la armonía entre deseos y

²⁰ Como hemos visto más arriba, Frankfurt considera falso el Principio de Posibilidades Alternativas.

acciones, y en el caso de Frankfurt, entre voluntad y voliciones de segundo orden, para Watson consiste en la adecuación entre valores y motivación.

Susan Wolf ha denominado las teorías anteriores “teorías del Yo Real”. Según estas teorías, la libertad o control que requiere la responsabilidad moral consiste en que la acción esté determinada por el yo real, cuyo núcleo, dependiendo de cada propuesta, está constituido por los deseos (de primer orden), las voliciones de segundo orden o el sistema axiológico de un agente. Para Wolf, el control sobre las propias acciones que proponen estas teorías no es suficientemente profundo para fundamentar la responsabilidad moral. Hay casos en que las acciones de una persona proceden de su yo real pero tenemos dudas fundadas sobre su responsabilidad, porque “tenemos a veces razones para cuestionar la responsabilidad de un agente *por* su yo real. Esto es, podemos pensar que no es culpa del agente ser la persona que es” (Wolf 1990, p. 37). Estos casos incluyen determinadas formas de enfermedad mental y de condicionamiento psicológico, así como casos de infancia fuertemente traumática. Esta crítica parece apuntar a la necesidad de una concepción más radical de la condición de origen o control último, de corte incompatibilista, según la cual el agente ha de ser responsable de sí mismo, de su propia estructura psicológica y axiológica, para ser responsable de las decisiones y acciones que resultan de ella. Sin embargo, Wolf considera incoherente esta versión radical del control. Según Wolf, la razón por la que, en las circunstancias mencionadas, cuestionamos la responsabilidad moral del agente no es la falta de ese control radical sobre su yo, sino más bien el hecho de que tales circunstancias privan al agente de la capacidad de formar creencias y valores correctos y coherentes entre sí. Así, según su propuesta, que ella denomina “la perspectiva de la Razón”, “un individuo es responsable si, y sólo si, es capaz de formar sus acciones sobre la base de sus valores [como en las teorías del Yo Real] y es capaz de formar sus valores sobre la base de lo que es Verdadero y Bueno” (Wolf 1990, p. 75).

Frente a las formas anteriores de control volitivo-racional, que pueden considerarse de carácter internista, en tanto que en ellas el control depende de hechos internos a la estructura psicológica de un agente y accesibles a su conciencia, Wolf introduce requisitos externistas, como son la verdad de sus creencias y la bondad de sus valores, a las que puede no tener acceso consciente.

Otros autores de orientación compatibilista han apuntado a otras condiciones de la responsabilidad moral, no necesariamente incompatibles con las anteriores. Así, por ejemplo, R. Jay Wallace (1994) considera que, para ser moralmente responsable de sus

acciones, un agente ha de tener la capacidad de aprehender razones morales y regular sus acciones según ellas. Y John Fischer ha puesto el acento en la capacidad de respuesta a razones (*reasons-responsiveness*), a la que ya hemos aludido.

Las condiciones anteriores constituyen especificaciones de ese aspecto de la libertad que hemos llamado control volitivo-racional. Es razonable pensar que estas condiciones, tal vez con algunas modificaciones, son necesarias para la responsabilidad moral. Pensemos, sin embargo, que un agente puede satisfacerlas debido a un proceso de manipulación neurológica. Si, en un caso así, negamos su responsabilidad moral, entonces dichas condiciones no son suficientes. Y, sobre este punto de partida, cabe construir un argumento general en contra del compatibilismo, que veremos en el siguiente apartado.

2. El Argumento de la Manipulación

Resulta esencial para el compatibilismo la posibilidad de trazar una distinción plausible entre la posesión y la carencia de responsabilidad moral (y por tanto de sus condiciones necesarias) sin abandonar el marco del determinismo. Éste es un supuesto de las propuestas presentadas en el apartado anterior. Para sus autores, y para el resto de compatibilistas, el determinismo como tal no afecta a la responsabilidad moral. Son otro tipo de factores (la compulsión, la falta de voliciones de segundo orden, la incapacidad de formar valores correctos o de responder a razones, etc.) los que establecen la diferencia entre agentes libres y moralmente responsables de sus acciones y agentes que carecen de esas propiedades, aun asumiendo la verdad del determinismo.

El llamado “Argumento de la Manipulación” trata de mostrar que el compatibilismo es falso, porque el determinismo elimina la responsabilidad moral. Según este argumento, si reflexionamos sobre las razones por las que no consideramos moralmente responsables a agentes cuya estructura psicológica y capacidades racionales han sido instauradas mediante la manipulación neurológica, esas mismas razones nos deberían llevar a un juicio semejante en el caso de que dicha estructura y capacidades se hayan desarrollado en un mundo determinista. La versión más detallada del argumento la debemos a Pereboom (cf. 2001, 2007, 2013).

Pereboom expone su argumento presentando cuatro casos que tienen en común una acción moralmente inaceptable (el asesinato de una colega, Ms. White) llevada a cabo por el protagonista, el profesor Plum. Un rasgo crucial, y también común a los cuatro casos, es el hecho de que, al decidir y actuar como lo hace, Plum satisface todas las

condiciones para la responsabilidad moral que hemos visto subrayadas por distintos autores compatibilistas en el apartado anterior. En los cuatro casos, esas condiciones, así como la acción misma que resulta de ellas, están causalmente determinadas; la diferencia reside en las distintas formas en que esa determinación causal tiene lugar.

En el caso 1, Plum ha sido creado por un grupo de neurocientíficos y su acción está causalmente determinada mediante la manipulación a distancia, por parte de ese grupo, de sus deseos, motivos y procesos de razonamiento y decisión. El argumento se basa en la intuición de que, en este caso, a pesar de que satisface todas las condiciones compatibilistas para la responsabilidad moral, Plum no es moralmente responsable de su acción; y la razón parece ser que su acción está causalmente determinada por los actos de los neurocientíficos, que se hallan más allá de su control.

En el caso 2, la determinación causal de la acción obedece a que los neurocientíficos han programado a Plum en los primeros días de su vida para que tenga las características y capacidades psicológicas que tiene y ejerce cuando decide matar a Ms. White, aunque ya no intervienen directamente en el proceso. Intuitivamente, Plum no es tampoco moralmente responsable, y de nuevo la razón parece ser la misma que en el caso 1, a saber, que su acción está causalmente determinada por factores sobre los que no tiene control. Las diferencias relativas al tipo de manipulación y al tiempo en que sucede no parecen relevantes para modificar el juicio negativo inicial sobre la responsabilidad moral de Plum.

En el caso 3, la determinación causal de la acción que Plum lleva a cabo deriva de la determinación causal de su carácter, rasgos y capacidades psicológicas a través de las rigurosas prácticas educativas de su familia y su comunidad social, que tienen lugar en una época demasiado temprana para que pueda eludirlas o modificarlas. Si el compatibilista pretende sostener que, en estas circunstancias, Plum es moralmente responsable, tendrá que aducir alguna diferencia relevante entre este caso y los dos anteriores. No podrá recurrir a una diferencia en la satisfacción de las condiciones compatibilistas de control volitivo-racional, que por estipulación se da en la misma medida en los tres casos. Dado que, en el caso anterior, la ausencia de responsabilidad moral se debe a la determinación causal de la acción de Plum por factores ajenos a su control, habrá que concluir, por la misma razón, que Plum no es tampoco responsable en este tercer caso.

Finalmente, en el caso 4, el determinismo causal es verdadero: todo lo que en el mundo sucede, incluyendo la decisión y la acción de Plum, es el resultado necesario e

inevitable del pasado y las leyes naturales. Plum es un ser humano normal, criado en circunstancias normales, y la determinación causal de su acción se produce solamente en virtud de la verdad del determinismo causal. Pero no parece haber diferencias relevantes entre este caso y el anterior, de modo que no hay justificación para sostener que en este último caso Plum es moralmente responsable si no lo era en el precedente. Apelar a que ahora la determinación causal no se lleva a cabo a manos de otros agentes no resulta convincente, ya que podríamos añadir un caso en el que los estados y procesos psicológicos de Plum son inducidos, no por un grupo de neurocientíficos, sino por un mecanismo generado espontáneamente, sin un diseñador inteligente. La mejor explicación del juicio sobre la carencia de responsabilidad de Plum en los tres primeros casos es la determinación causal de su acción por factores que escapan a su control. Y, puesto que ésta es también la situación en el último caso, el juicio ha de ser también el mismo: Plum no es tampoco moralmente responsable en este caso.

Así, si el argumento es correcto, el determinismo causal como tal excluye la responsabilidad moral y el compatibilismo es falso. Las condiciones de control volitivo-racional propuestas por los compatibilistas pueden ser tal vez necesarias, pero no son suficientes para la responsabilidad moral. Como lo sugieren nuestros juicios sobre los casos anteriores, la responsabilidad moral por las acciones requiere que el agente sea el origen genuino de las mismas, lo que supone un control sobre los factores que le llevan a decidir y a actuar que el determinismo hace imposible. La condición de origen y control último no puede ser satisfecha asumiendo el determinismo, y los intentos compatibilistas de integrar en sus propuestas las intuiciones que subyacen a esa condición quedan muy lejos de tener éxito.

3. Respuestas compatibilistas al argumento de la manipulación

El compatibilismo, sin embargo, no está indefenso frente al argumento de la manipulación. Podemos distinguir dos tipos de respuesta que el compatibilista puede dar a este argumento: de línea dura y de línea blanda²¹. El núcleo de la primera consiste en desactivar el argumento en su base atribuyendo al agente responsabilidad moral en (prácticamente) todos los casos considerados. Podemos tomar a Frankfurt como representante de esta estrategia. Así, en un primer momento (1975, p. 53), Frankfurt sostuvo que, si la manipulación toma la forma de una programación (lo que

²¹ Tomo esta terminología de Patarroyo 2010, cap. III. Esta excelente tesis doctoral me ha sido de gran ayuda en la elaboración de algunas partes del presente trabajo.

correspondería al caso 2 del argumento de Pereboom) que dota “de un carácter o programa estable” al sujeto, éste “puede llegar a ser moralmente responsable, suponiendo que esté adecuadamente programado, del mismo modo que otros: identificándose con algunos de sus deseos de segundo orden” (1975, p. 53) y asumiendo la responsabilidad por ellos y por las acciones a las que le conducen. En ese momento, en cambio, Frankfurt sostiene que, si la manipulación tiene lugar “sobre una base continua ... de modo que cada estado mental del sujeto es el resultado de una intervención específica” (1975, p. 53), lo que corresponde al caso 1 del argumento de Pereboom, dicho sujeto no es moralmente responsable; pero la razón no es la falta de control u origen último, sino más bien que “el sujeto no es una persona en absoluto. Su historia es completamente episódica y carente de conexión interna” (1975, p. 53). Sin embargo, como Pereboom objeta, un sujeto manipulado sobre una base continua puede tener una vida mental plenamente conexa e internamente indistinguible de la de una persona no manipulada. De hecho, ésta es la situación en el caso 1 del argumento de Pereboom. Pues bien, más adelante Frankfurt responderá que, bajo ese supuesto, el sujeto puede ser también moralmente responsable en el caso 1:

A través de su intervención, un manipulador puede tener éxito en proveer a una persona, no sólo de sentimientos y pensamientos particulares, sino también de un nuevo carácter. La persona es entonces moralmente responsable de las elecciones y la conducta a las que éste carácter le lleva. Después de todo, estamos inevitablemente configurados y sustentados por circunstancias sobre las que no tenemos control alguno. Las causas a las que estamos sujetos pueden también cambiarnos radicalmente, sin por ello hacer que no seamos agentes moralmente responsables. Es irrelevante que esas causas operen en virtud de fuerzas naturales que configuran nuestro entorno o mediante los designios deliberadamente manipuladores de otros agentes humanos. Somos la clase de personas que somos; y lo que cuenta es lo que somos, en lugar de la historia de nuestro desarrollo. (Frankfurt 2002, p. 28)

Lo importante para la responsabilidad moral, según Frankfurt, es que la persona posea un carácter y una estructura psíquica con un grado de complejidad suficiente para tener voliciones de segundo orden e identificarse, a través de ellas, con los deseos que le llevan a actuar. No importa cuál es el origen de ese carácter y esa estructura psíquica. La

concepción de la responsabilidad moral en Frankfurt es ahistórica. Y éste es también el caso por lo que se refiere a Watson y Wolf. Así, en coherencia con su posición, y suponiendo, plausiblemente, que, en los cuatro casos del argumento de la manipulación, el agente puede ser movido a actuar por sus valores (Watson) y formar valores correctos y coherentes (Wolf), estos autores habrían de seguir también la línea dura de Frankfurt y aceptar que Plum es moralmente responsable en (prácticamente) todos los casos imaginados por Pereboom.

Sin embargo, resulta difícil admitir que una persona a la que se ha proporcionado un nuevo carácter mediante un proceso de manipulación psíquica o neurológica sea moralmente responsable de las acciones que derivan de dicho carácter, en el sentido de merecer realmente reprobación o castigo por ellas. Así, algunos compatibilistas, como Fischer y Ravizza o, más recientemente, Haji y Cuypers, han aceptado el punto de partida del argumento, negando la responsabilidad moral del agente en casos de manipulación pero afirmándola cuando ésta no se da, lo que sucede sobre todo en el caso 4 del argumento. Con ello, adquieren el compromiso de justificar esta diferencia de juicio apelando a la diferencia en la forma específica que adopta la determinación causal de la estructura psíquica, capacidades, decisiones y acciones del sujeto en los distintos casos. De aquí deriva una respuesta de línea blanda al argumento de la manipulación.

Para Fischer y Ravizza, además de cumplir las condiciones de racionalidad y respuesta a razones a las que nos hemos referido, el mecanismo de deliberación ha de pertenecer realmente al agente, lo cual requiere un determinado tipo de origen: el agente ha de ser entrenado por sus padres y educadores en el proceso de asumir responsabilidad por sus acciones, un proceso a través del cual va formando creencias acerca de su propia condición de objeto apropiado de aprobación o reprobación y acaba convirtiéndose en un agente moralmente responsable (cf. Fischer y Ravizza 1998, p. 210). Este no es el origen de la estructura psíquica del agente en los casos de manipulación, lo que da cuenta del juicio según el cual, en ellos, a diferencia del caso puramente determinista, no tenemos un agente moralmente responsable.

Por lo que respecta a Cuypers y Haji, es central la distinción entre vías normales y vías causalmente desviadas por las que un agente adquiere factores psicológicos relevantes para la responsabilidad moral: “En los casos problemáticos de manipulación, elementos psicológicos tales como deseos o creencias, entre otras cosas, se adquieren a través de vías causales que son desviadas en comparación con vías causales consideradas normales” (Haji y Cuypers 2008, p. 15; cit. en Patarroyo 2010, p. 276). Dadas ciertas

condiciones, los factores adquiridos por estas últimas vías, y en particular el esquema evaluativo del agente, pertenecen realmente al agente y sustentan su responsabilidad moral, a diferencia de los obtenidos a través de vías desviadas.

Estas propuestas tratan de mostrar que las intuiciones que subyacen a la condición de control u origen último para la responsabilidad moral pueden ser satisfechas en un marco determinista. No parecen, sin embargo, capaces de resolver los problemas de suficiencia que afectan al compatibilismo. Tanto Fischer y Ravizza como Haji y Cuypers parecen confundir la descripción de nuestras prácticas vinculadas a la responsabilidad moral y de las distinciones internas a las mismas con la justificación de esas prácticas y esas distinciones. Es cierto que atribuimos responsabilidad moral a sujetos que se han desarrollado en ciertas condiciones de normalidad y no a otros que han sido manipulados; el problema, sin embargo, es la justificación racional de esa diferencia *una vez que se presupone el determinismo*. Un incompatibilista no tiene por qué negar que, aun cuando el determinismo sea verdadero, hay muchas diferencias humanamente importantes entre un contexto evolutivo normal, que promueva el desarrollo de cualidades intelectuales y morales, y un contexto de manipulación; lo que niega es que haya una diferencia en la responsabilidad moral. Como Pereboom señala, si la razón por la que negamos responsabilidad moral a determinados agentes sometidos a manipulación es su falta de control sobre los factores que dan lugar a sus acciones, esa misma falta de control tiene lugar si el determinismo es verdadero, por lo que el juicio sobre la responsabilidad moral debería ser asimismo negativo.

La lección que de todo esto parece desprenderse es la necesidad, para la responsabilidad moral, del origen o control último interpretado de un modo incompatible con el determinismo. Los intentos compatibilistas de interpretar esta condición en consonancia con el determinismo dan lugar a sucedáneos insatisfactorios. El problema, sin embargo, como distintos autores han sostenido, es que la satisfacción de esta condición parece también imposible si el determinismo es falso. Y, bajo determinadas interpretaciones del contenido de dicha condición, no podemos sino darles la razón, como veremos a continuación.

V. EL PROBLEMA DEL CONTROL ÚLTIMO Y EL LIBERTARISMO

En la interpretación de Nietzsche, lo que nosotros llamamos “control último” y él denomina “responsabilidad total y última” de las propias acciones o, simplemente, “libertad de la voluntad” o “voluntad libre”, involucra la imposibilidad de ser *causa sui*:

La *causa sui* [causa de sí mismo] es la mejor autocontradicción imaginada hasta ahora, una especie de estupro y monstruosidad lógicos: pero el desenfrenado orgullo del hombre le ha llevado a enredarse de manera profunda y horrible justo en ese sinsentido. La aspiración a la “libertad de la voluntad”, entendida en aquel sentido metafísico y superlativo que, por desgracia, continúa dominando en las cabezas de los semiinstruidos, la aspiración a cargar uno mismo con la responsabilidad total y última de sus propias acciones, y a descargar de ella a Dios, al mundo, a los antepasados, al azar, a la sociedad, equivale, en efecto, nada menos que a ser precisamente aquella *causa sui* y a sacarse a sí mismo de la ciénaga de la nada y a salir a la existencia a base de tirarse de los cabellos, con una temeridad mayor aún que la de Münchhausen. (Nietzsche 1886/1972, p. 42)

Más recientemente, Galen Strawson (1986) ha sostenido igualmente que el control o la libertad que serían necesarios para la responsabilidad moral son lógicamente imposibles, ya que requieren que se complete efectivamente “un regreso infinito de elecciones de principios de elección” (Strawson 1986, p. 29). Según Strawson, puesto que el modo como actuamos depende del modo como somos, hemos de ser responsables del modo como somos para ser responsables del modo como actuamos. Ésta es la idea del control último en su versión más radical. Ahora bien, según Strawson, para ser responsables del modo como somos hemos de haber elegido, consciente y racionalmente, ser como somos, para lo cual hemos de haber hecho esa elección sobre la base de algunos criterios o principios. Pero, puesto que esta elección de nuestro carácter y forma de ser depende de dichos criterios o principios, hemos de haber elegido consciente y racionalmente esos criterios o principios para ser responsables de esa elección de nuestro carácter y forma de ser, y con ello de estos últimos. Para ello, sin embargo, hemos de existir ya dotados con otro conjunto de criterios de elección a la luz de los cuales elegir esos principios o criterios con los que elegimos cómo ser, y así sucesivamente. Hemos entrado, pues, en un regreso infinito de elecciones de criterios de elección que no podemos completar. La conclusión del argumento es la imposibilidad lógica de la “verdadera autodeterminación” o “verdadera responsabilidad”, en palabras de Strawson, de la “responsabilidad total y última” de nuestras acciones, en palabras de Nietzsche, o del control u origen último, en nuestra terminología. De hecho, en una formulación más breve de su argumento, Strawson incluye la premisa según la cual “para ser de verdad

moralmente responsable de las propias acciones, un agente tendría que ser *causa sui*, al menos en ciertos aspectos mentales cruciales” (Strawson 1994, p. 212).

Desde luego, si Nietzsche y Strawson tienen razón sobre lo que implica el control último sobre las propias acciones, y si el control último así entendido es necesario para la responsabilidad moral, nadie puede ser moralmente responsable de sus acciones. Por mi parte, sin embargo, aunque acepto la necesidad del control último para la responsabilidad moral, en el sentido de que hemos de ser responsables de (al menos algunos de) los factores y criterios sobre cuya base decidimos y actuamos para ser responsables de esas decisiones y acciones, rechazo que ello implique ser causa de sí mismo o un regreso infinito de elecciones de criterios de elección. Cualquier concepción razonable de la libertad y la responsabilidad moral ha de aceptar, como punto de partida, que no tenemos control ni elección alguna sobre los factores que configuran inicialmente nuestra personalidad. Así, nuestras elecciones iniciales estarán necesariamente condicionadas por dichos factores. Pero condicionamiento no significa determinación. Nuestra dotación psicológica inicial, suponiendo ciertas condiciones de normalidad en nuestras capacidades mentales y en nuestro entorno, limita el rango de nuestras elecciones, pero no lo reduce a cero.

En realidad, una elección absolutamente incondicionada es una elección sin criterios, y con ello una elección puramente arbitraria, de la que el sujeto no es moralmente responsable. Pero, frente a Strawson, no necesitamos haber *elegido* nuestros criterios de elección para ser responsables de la elección misma, del mismo modo que no necesitamos haber elegido las reglas del ajedrez o ciertos principios de estrategia ajedrecística para ser positivamente responsables de haber realizado una buena jugada. Y, paralelamente, no hemos de haber elegido ni creado nuestros principios y criterios morales para ser responsables de las decisiones que tomamos basándonos en ellos. Pero no estamos tampoco obligados a aceptarlos ciegamente; una aceptación ciega erosiona asimismo el control último y la responsabilidad moral. Podemos hacer nuestros esos criterios o principios formando creencias justificadas sobre su corrección o verdad, una justificación que puede incluir argumentos de autoridad basados en la admiración y el respeto hacia determinadas personas, pero que no se reduce a ellos. En realidad, al adquirir creencias en el curso de nuestro desarrollo personal, adquirimos también criterios para cuestionarlas y modificarlas, tales como la coherencia con otras creencias, la aspiración a la verdad o la adecuación a la experiencia. En ausencia de estos criterios, no cabría hablar de un sistema racional de creencias. Y ello se aplica asimismo a nuestras

creencias valorativas, que incluyen nuestros valores morales. En Moya 2006 (cap. 5) he defendido la existencia de una forma de control sobre nuestras creencias, incluyendo en especial nuestras creencias valorativas (que subyacen centralmente a nuestras decisiones moralmente relevantes) que no descansa en elecciones o actos de voluntad, sino en una actitud de respeto y receptividad hacia la verdad y lo valioso, así como a los demás criterios de justificación y revisión de creencias. Esta actitud conlleva también el abandono de algunas de nuestras convicciones y la adopción de otras. En este proceso vamos formando un sistema de creencias que realmente podemos llamar nuestro. Y este control no volitivo sobre nuestras creencias valorativas nos proporciona también control sobre las decisiones que tomamos basándonos en ellas, haciéndonos así autores genuinos de las mismas. Esta autoría u origen último es posible porque la aplicación de criterios generales para la formación de creencias valorativas y, a su vez, la aplicación de estas últimas a situaciones concretas conlleva un grado importante de indeterminación, lo que confiere a nuestra resolución en cada caso un papel genuinamente originante. Así, por ejemplo, aun cuando la amistad forme parte de nuestro sistema de valores, en muchas situaciones concretas no está nada claro en qué consistiría hacer honor a ese valor; y resolver esta cuestión requiere cosas tales como la aplicación reflexiva de criterios, la capacidad de ponerse en lugar del otro, el recuerdo de experiencias previas, etc. El resultado de este proceso no está determinado por todos estos factores; pero no es tampoco arbitrario o irracional, sino, al menos en parte, un fruto de la reflexión y el razonamiento práctico.

En mi opinión, esta aproximación cognitiva a la libertad y la responsabilidad moral puede combinarse fructíferamente con la concepción volitiva de Robert Kane sobre el control o responsabilidad última, que bosquejaremos a continuación, para dar lugar a una perspectiva razonable, de inspiración libertarista, sobre la libertad y la responsabilidad moral. Esta perspectiva podría acabar de atajar una objeción tradicional al libertarismo a la que ya me he referido en la sección I y que ya he comenzado a responder en el párrafo anterior. Según esta objeción, que cabe denominar la objeción de la arbitrariedad, la indeterminación causal de las decisiones hace de ellas actos arbitrarios y azarosos, que no satisfacen la condición de control volitivo-racional por parte del agente, y de los cuales no puede éste ser responsable.

En la concepción de Kane (cf. 1996, 2007), el control o responsabilidad última descansa en ciertas elecciones que él denomina “autoformadoras”, en el sentido de que contribuyen a configurar nuestro propio carácter y a hacernos responsables del mismo y

de las acciones que ulteriormente puedan brotar de él. Estas elecciones se plantean a agentes dotados ya de un mínimo de capacidades racionales y sensibilidad moral; no parten de la nada. Para cumplir su función de autoformación, estas elecciones han de satisfacer ciertos requisitos. En primer lugar, el agente ha de tener conjuntos de razones *inconmensurables* (por ejemplo, razones morales y razones de interés propio) para cada una de las opciones que considera. En segundo lugar, ninguno de esos conjuntos de razones ha de ser causalmente suficiente para la elección final, de modo que ésta se halla causalmente *indeterminada*. Esto no significa que la elección no esté causada, al menos parcialmente, por esas razones; sólo supone que la relación causal no es determinista, sino probabilística. Pero la indeterminación causal de la acción por cada conjunto de razones resulta también de la presencia del otro conjunto, que inclina la voluntad en sentido opuesto al primero. En tercer lugar, y como consecuencia de lo anterior, la elección satisface la condición de *posibilidades alternativas*: sea cual sea la decisión del agente, éste hubiera podido tomar, en sentido categórico, una decisión distinta. En estas condiciones, la decisión final, cualquiera que sea, resultará del esfuerzo del agente por superar el conflicto y se llevará a cabo de manera libre, voluntaria y racional: no será una decisión arbitraria. Y al mismo tiempo contribuirá a conformar la voluntad y la forma de ser y de vivir del agente. Es razonable interpretar que elecciones de este tipo se llevan a cabo a lo largo de toda la vida de una persona, de modo que ésta nunca se halla completamente formada. Así, gracias a estas elecciones autoformadoras y a las correspondientes acciones, cuando un agente tiene razones decisivas para decidir y actuar de cierto modo, de forma que una decisión y una acción alternativa sería incomprensible o irracional, esa decisión se halla bajo el control último del agente, pues resulta de una personalidad que él mismo ha contribuido a formar y con respecto a la cual tuvo alternativas.

Sin embargo, una concepción de la responsabilidad moral como la de Kane, que hace descansar el control sobre las propias acciones en elecciones o actos de voluntad, se enfrenta a dificultades. En su concepción de las elecciones autoformadoras, aun cuando las razones no determinen causalmente la decisión, no dejan por ello de causarla (de modo probabilístico). Así, la decisión es, en una medida importante, función de esas razones, de modo que, siguiendo la lógica del control último, el agente debería tener control sobre esas razones para tener control sobre la decisión que finalmente toma. Pero si el control sobre esas razones consiste también en una elección de las mismas, estamos iniciando, al parecer, un regreso infinito no muy lejano al que Strawson apunta en su argumento. Este

regreso sólo se detendría en una elección absolutamente incondicionada, carente por completo de razones y por ello totalmente arbitraria. Una elección así no sería libre, ya que violaría la condición de control volitivo-racional, y no podría fundar la responsabilidad moral. Es en este punto donde el concepto de un control no basado en elecciones o actos de voluntad, que forma parte de la aproximación cognitiva a la libertad y la responsabilidad moral que propugno, puede desempeñar un papel importante en la defensa de una concepción libertarista, que incluya la condición de control último, frente a la objeción de la arbitrariedad. En efecto, el control del agente sobre las razones presentes en las situaciones de elección autoformadora no tiene por qué consistir en una elección de esas razones; en la medida en que éstas consisten, en gran medida, en creencias valorativas, el control sobre ellas, que las convierte en parte del agente mismo, puede basarse en una actitud pasiva de respeto y reconocimiento de su verdad, corrección o justificación, en el sentido más arriba indicado. Esta forma no volitiva de control permite evitar el regreso infinito de elecciones de criterios de elección o su detención mediante una elección absolutamente arbitraria, preservando el control racional del agente sobre sus criterios de elección. Tendríamos con ello, frente a Nietzsche y Strawson, una concepción razonable, al alcance de seres humanos finitos como nosotros, del control último como requisito de la responsabilidad moral. Pero un desarrollo más detallado de esta propuesta desbordaría los límites de este ya extenso trabajo.

CONCLUSIÓN

Este trabajo comenzó con un análisis de los conceptos de determinismo, libertad y responsabilidad moral. En relación con la libertad, he distinguido tres aspectos o condiciones de la misma: posibilidades alternativas, control volitivo-racional y control u origen último. Aunque estas tres condiciones parecen inicialmente necesarias para la responsabilidad moral, no todos aceptan que realmente lo son; y hay también divergencias por lo que respecta a su compatibilidad con el determinismo. La actitud ante estas dos cuestiones define las principales posiciones en el debate actual sobre las relaciones entre libertad, responsabilidad moral y determinismo: por un lado, el incompatibilismo en distintas variedades, como el libertarismo, el determinismo estricto y el incompatibilismo estricto y, por el otro, el compatibilismo y su variedad semicompatibilista. He argüido a favor de la incompatibilidad del determinismo con las posibilidades alternativas y el control último. Y he defendido la necesidad de las alternativas y el control último para la responsabilidad moral frente a posiciones que niegan dicha necesidad. Con respecto al

control último, he argumentado además a favor de su coherencia y posibilidad, frente a perspectivas que las rechazan. He sostenido, pues, una posición general de carácter incompatibilista y libertarista. Por lo que respecta al control volitivo-racional, he defendido, frente a argumentos como el de la arbitrariedad, su viabilidad en un marco indeterminista. Finalmente, he sugerido la posibilidad de unir, con importantes ventajas teóricas, la concepción volitiva de la responsabilidad moral de Robert Kane con la concepción cognitiva que yo defiendo.

Los argumentos y consideraciones que he aportado a favor de las tesis indicadas están muy lejos de ser definitivos. El debate sobre la libertad, la responsabilidad moral y el determinismo no tiene visos de llegar, en un futuro próximo, a una conclusión satisfactoria para todos los participantes en el mismo.

REFERENCIAS BIBLIOGRÁFICAS

- Austin, J. L. (1970), *Philosophical Papers*, Oxford, Oxford University Press.
- Berofsky, B. (2009), "Free will and the mind-body problem", *Australasian Journal of Philosophy* 88, pp. 1-19.
- Chisholm, Roderick (1964), "Human freedom and the self", The Lindley Lecture, University of Kansas; reimpresso en Watson (ed.) 2003, pp. 26-37.
- Davidson, Donald (1982), "Freedom to act", en *Essays on Actions and Events*, Oxford, Clarendon Press (tr. cast.: *Ensayos sobre acciones y sucesos*, Barcelona, Crítica/UNAM), pp. 63-81.
- Fischer, John M. (1994), *The Metaphysics of Free Will*, Oxford, Blackwell.
- (2007), "Compatibilism", en J. M. Fischer, R. Kane, D. Pereboom y M. Vargas, *Four Views on Free Will*, Oxford, Blackwell Publishing, pp. 44-84.
- y Ravizza, Mark (1998), *Responsibility and Control: a theory of moral responsibility*, Cambridge, Cambridge University Press.

- Frankfurt, Harry G. (1969), "Alternate possibilities and moral responsibility", *Journal of Philosophy* 66, pp. 829-839. Reimpreso (entre otros lugares) en Watson, ed., 2003, pp. 167-176.
- (1971), "Freedom of the will and the concept of a person", *Journal of Philosophy* 68, pp. 5-20; reimpreso en (y citado según) *The Importance of What We Care About*, Cambridge, Cambridge University Press, 1988, pp. 11-25.
- (1975), "Three concepts of free action", *Proceedings of the Aristotelian Society*, Suppl. Vol. 49; reimpreso en (y citado según) *The Importance of What We Care About*, Cambridge, Cambridge University Press, 1988, pp. 47-57.
- (2002), "Reply to John Martin Fischer", en Sarah Buss y Lee Overton (eds.), *Contours of Agency. Essays on themes from Harry Frankfurt*, Cambridge MA, MIT Press, pp. 27-31.
- Ginet, Carl (1996), "In defense of the principle of alternative possibilities: why I don't find Frankfurt's argument convincing", *Philosophical Perspectives* 10, pp. 403-417.
- Haji, Ishtiyaque y Cuypers, Stefaan (2008), *Moral Responsibility, Authenticity and Education*, Nueva York, Routledge.
- Hume, David (1748/2004), *Investigación sobre el entendimiento humano* (tr. V. Sanfélix y C. Ors), Madrid, Istmo.
- Kane, Robert (1985), *Free Will and Values*, Albany, NY, State University of New York Press.
- (1996), *The Significance of Free Will*, Oxford y Nueva York, Oxford University Press.
- (2007), "Libertarianism", en J. M. Fischer, R. Kane, D. Pereboom y M. Vargas, *Four Views on Free Will*, Oxford, Blackwell Publishing, pp. 5-43.
- (2009), "Free Will and the Dialectic of Selfhood: Can We Make Sense of a Traditional Free Will Requiring Ultimate Responsibility?", *Ideas y Valores* 58, pp. 25-43.

Lewis, David (1981), “Are we free to break the laws?”, *Theoria* 47, pp. 112-21. Reimpreso en Watson, ed., 2003, pp. 122-129.

Moya, Carlos J. (2003), “Blockage cases: no case against PAP”, *Crítica* 35, pp. 109-120.

—— (2006), *Moral Responsibility: the ways of scepticism*, Abingdon y Nueva York, Routledge.

—— (2007), “Moral responsibility without alternative possibilities?”, *Journal of Philosophy* 104, pp. 475-486.

—— (2009), “Alternativas, responsabilidad y respuesta a razones”, *Ideas y Valores* 58, pp. 45-65.

—— (2010), “Blame, moral obligation and alternative possibilities”, manuscrito, presentado como ponencia inaugural en el VI Congreso de la Sociedad Española de Filosofía Analítica, Tenerife.

—— (2011), “On the very idea of a robust alternative”, *Crítica* 43, pp. 3-26.

—— (2014), “Doing one’s best, alternative possibilities, and blameworthiness”, *Crítica* 46, pp. 3-26.

—— (en prensa), “Alternatives and responsibility: an asymmetrical approach”, en Anne Reboul (ed.), *Mind, Values, and Metaphysics*, Springer.

Naylor, Margery B. (1984), “Frankfurt on the principle of alternate possibilities”, *Philosophical Studies* 46, pp. 249-258.

Nietzsche, Friedrich (1886/1972), *Más allá del bien y del mal*, tr. de A. Sánchez Pascual, Madrid, Alianza Editorial.

- Patarroyo, Carlos G. (2010), *Un lugar para la responsabilidad moral. Una defensa del semi-compatibilismo*. Tesis doctoral (Luis E. Hoyos, director). Universidad Nacional de Colombia, Facultad de Ciencias Humanas, Departamento de Filosofía.
- Pereboom, Derk (2001), *Living Without Free Will*, Nueva York, Cambridge University Press.
- (2007), “Hard Incompatibilism”, en J. M. Fischer, R. Kane, D. Pereboom y M. Vargas, *Four Views on Free Will*, Oxford, Blackwell Publishing, pp. 85-125.
- (2009), “Further thoughts about a Frankfurt-style argument”, *Philosophical Explorations* 12, pp. 109-118.
- (2012), “Frankfurt cases, derivative responsibility, and the timing objection”, *Philosophical Issues* 22, pp. 298-315.
- (2013), “Optimistic Skepticism about Free Will”, en P. Russell y O. Deery (eds.), *The Philosophy of Free Will. Selected contemporary readings*, New York, Oxford University Press, pp. 421-449.
- Pink, Thomas (1996), *The Psychology of Freedom*, Cambridge, Cambridge University Press.
- Scanlon, Thomas M. (1988), “The significance of choice”, en S. M. McMurrin (ed.), *The Tanner Lectures on Human Values*, University of Utah Press, pp. 151-177; reimpresso en Watson, ed. (2003), pp. 352-371.
- (1998), *What We Owe to Each Other*, Cambridge MA, Harvard University Press.
- Strawson, Galen (1986), *Freedom and Belief*, Oxford, Clarendon Press.
- (1994), “The impossibility of moral responsibility”, *Philosophical Studies* 75, pp. 5-24; reimpresso en (y citado según) Watson (ed.) 2003, pp. 212-228.

Strawson, Peter F. (1962), "Freedom and resentment", *Proceedings of the British Academy* 48, pp. 1-25; reimpresso en Watson (ed.) 2003, pp. 72-93. (Tr. cast. de J. J. Acero en Strawson, *Libertad y resentimiento*, Barcelona, Paidós/UAB, pp. 37-67).

Van Inwagen, Peter (1983), *An Essay on Free Will*, Oxford, Clarendon Press.

Unger, Peter (2002), "Free will and scientificism", *Philosophy and Phenomenological Research* 65, pp. 1-25.

Wallace, R. Jay (1994), *Responsibility and the Moral Sentiments*, Cambridge MA, Harvard University Press.

Warfield, Ted A. (2000), "Causal determinism and human freedom are incompatible: a new argument for incompatibilism", *Philosophical Perspectives* 14, pp. 167-180.

Watson, Gary (1987), "Free action and free will", *Mind* 96, pp. 145-172.

—— (ed.) (2003), *Free Will* (2^a ed.), Oxford, Oxford University Press.

Widerker, David (1995), "Libertarianism and Frankfurt's attack on the principle of alternative possibilities", *Philosophical Review* 104, pp. 247-261.

—— (2006), "Libertarianism and the philosophical significance of Frankfurt scenarios", *Journal of Philosophy* 103, pp. 169-187.

Wolf, Susan (1990), *Freedom Within Reason*, Oxford, Oxford University Press