

Voluntad y responsabilidad moral

(Voluntariness and Moral Responsibility)

Sergi ROSELL

Recibido: 24 de septiembre de 2012

Aceptado: 16 de abril de 2013

Resumen

En este artículo, me opongo a la idea de que toda acción debe ser moralmente evaluada *exclusivamente* por el motivo o intención por el que el agente la realizó o intentó realizarla. En particular, se presentan casos con los que se discute el contraste entre voluntad o intención y consecuencias; entre negligencia y decisiones en condiciones de incertidumbre, y resultados; entre actos deliberados e inadvertidos; y entre intención y acción; con la meta de mostrar el papel adicional que el segundo elemento de cada par desempeña en la determinación del juicio moral merecido.

Palabras clave: responsabilidad moral, voluntad, intención, motivo, acción, negligencia, incertidumbre, consecuencias.

Abstract

In this article I argue against the idea that actions are to be morally judged *only* for the motive or intention out of which the agent performed or intended to perform the action. Particularly, I put forward different cases by which I discuss the contrast between will or intention and consequences; between negligence and decisions under uncertainty, and outcomes; between deliberate and inadvertent acts; and between intention and action; aiming to show the additional role played by each pair's second element in determining the deserved moral judgment.

Keywords: Moral Responsibility, Will, Intention, Motive, Action, Negligence, Uncertainty, Consequences.

1. Introducción

Para cierto sentido común, así como para algunos insignes filósofos, nuestros actos deben juzgarse exclusivamente sobre la base del motivo o intención con que el agente realizó el acto en cuestión. Adam Smith, en su *Teoría de los sentimientos morales*, afirma que la siguiente máxima es “reconocida por todo el mundo, y no hay ninguna voz que disienta entre toda la humanidad”:

Para su propiedad o impropiedad, para su beneficencia o carácter dañino [...] *todo elogio y censura, toda aprobación o desaprobación*, de cualquier tipo, que puede justificadamente ser emitida sobre cualquier acción, *debe últimamente pertenecer a la intención o afectación del corazón*. (II, § iii; cursivas añadidas)¹

Esta máxima puede justificarse o motivarse a partir de la idea contenida en las conocidas palabras con las que Kant inicia la *Fundamentación*:

La buena voluntad no es buena por lo que efectúe o realice, no es buena por su adecuación para alcanzar algún fin que nos hayamos propuesto; *es buena sólo por el querer, es decir, es buena en sí misma*. [...] *La utilidad o la esterilidad no pueden ni añadir ni quitar nada a ese valor*. (Parte I; Ak. IV, 411; mis cursivas)

Puede ser útil distinguir aquí entre motivo e intención, como hace Bernard Williams:

Del mismo modo que en el reino del *carácter* lo que cuenta es el *motivo*, y no la manera de ser, las potencias o la dotación; asimismo, en la *acción* no son los cambios efectivamente realizados en el mundo lo que cuenta, sino *la intención*. (Williams 1981, p. 20; mis cursivas)

En todo caso, lo fundamental es que para esta posición el valor o merecimiento del agente debería medirse o determinarse, exclusivamente, en función del motivo o intención por el que actúa; o, en general, de su voluntad, o del *signo* de su voluntad: buena o mala, y cómo de buena o mala. Llamemos a esto *el primado de la voluntad*:

(PV) Toda acción debe ser moralmente evaluada *exclusivamente* en relación al signo de la voluntad (motivo, intención) por la que el agente lleva a cabo (o intenta llevar a cabo) esta acción.

¹ Si bien la posición de Smith es luego más compleja. Véase Russell (1999) para un estudio de los pormenores de su posición.

Este principio se opone a que factores ulteriores –o, como algunos dirían, *contingentes*–, sean estos afortunados o desafortunados, puedan interferir en la evaluación moral que el agente merece.

En este artículo, presentaré diferentes tipos de casos que parecen contradecir los presupuestos de esta posición (PV), para tratar de cuestionarla. En concreto, en la sección primera me ocuparé del contraste entre voluntad o intención y las consecuencias de nuestras acciones; mientras que en la segunda me centraré en casos de negligencia y decisión en condiciones de incertidumbre. En la sección tercera me ocuparé de la distinción entre actos realizados deliberadamente y actos realizados inadvertidamente, y del papel que esta distinción debe desempeñar en relación a los límites de nuestra responsabilidad moral. Finalmente, en la sección cuarta, trataré de cuestionar que pueda establecerse una distinción nítida entre intención y acción, como asume el defensor del *primado de la voluntad*. Mi objetivo no es negar la importancia de la voluntad (intención, motivo) en la evaluación moral de nuestras acciones, sino la creencia de que es lo *único* que debe tenerse en cuenta.²

Cabe añadir que el tipo de evaluación moral que aquí nos interesa es el paradigmáticamente representado por juicios de responsabilidad moral que responden al merecimiento moral –que algunos llaman *genuino* o *último*– del agente. Por supuesto, estos juicios de responsabilidad moral no deben identificarse con juicios de responsabilidad legal. (Ciertamente, las relaciones entre moral y derecho son complicadas; pero, en todo caso, es claro que es posible que alguien sea moralmente responsable por algo, aunque no lo sea legalmente, y viceversa.)

2. Voluntad, intención e intento

Consideremos el siguiente caso.³ Dos personas tratan, independientemente, de asesinar a otras dos, llevando a cabo todos los preparativos necesarios para ello. Intencional y deliberadamente, ambos tratan de acabar con la vida de sus respectivas víctimas. Pero, finalmente, solo uno logra su objetivo y mata a su víctima, pues el disparo del otro resulta no ser mortal. El defensor del primado de la voluntad (PV) se opondrá aquí a que el mero carácter certero (mortal) o no del disparo pueda mar-

² Aunque en este artículo prefiero abordar la cuestión en estos términos, sin duda el rechazo del primado de la voluntad supondría una contribución importante a la vindicación de la *suerte moral consecucional* o *resultante* –esto es, la suerte moral debida a la realización y a los resultados o consecuencias de nuestras acciones. Sobre la cuestión de la suerte moral, en general, puede verse Rosell (2006), o mi introducción a Williams y Nagel (2013). Para un argumento contra la negación global de la suerte moral, véase Rosell (2012).

³ Muchos de los ejemplos propuestos a lo largo del artículo tienen su origen en Nagel (1979), si bien han sido posteriormente usados y reelaborados por otros autores y aquí aparecerán modificados según convenga. Otros son originales míos.

car una diferencia en el juicio moral que uno y otro merecen, y situará el *locus* de la atribución de responsabilidad moral en la volición e intención previa de los agentes. Lo mismo sucedería en el caso de que, en el último instante, uno de ellos no hubiera podido disparar a causa, por ejemplo, de la presencia de un obstáculo. En ambos casos, uno y otro serán igualmente responsables y culpables, en tanto que ambos intentaron igualmente acabar con la vida de una persona. De este modo, la diferencia legal entre homicidio e intento de homicidio sería *moralmente* irrelevante.

No obstante, parece obvio que uno es responsable de matar *a alguien*, mientras que el otro no puede ser responsable *de* matar a nadie, pues no hay nadie que haya muerto a causa de su disparo. Uno es responsable de tratar de matar a alguien y lograrlo, mientras que el otro lo es solo del intento o tentativa. Por lo tanto, hay un sentido claro en el que uno es más responsable que el otro. Pero el defensor del primado de la voluntad puede responder de este modo:

No niego en absoluto que [uno] sea *más responsable* que [el otro]. Debemos distinguir el *grado* en la responsabilidad de alguien, de su *alcance*. [...] Mi tesis es que [uno] y [otro] cargan con el mismo grado de responsabilidad, más allá del hecho de que la responsabilidad de [uno] tenga un alcance mayor. (Zimmerman 2002, p. 560-1)

Aunque uno sea responsable *de más cosas* que el otro, ambos son responsables *en el mismo grado*.⁴ Así, según esta posición, ambos son igualmente responsables de su intención e intento de acabar con la vida de una persona, pues “[e]l hecho de que el intento de [uno] no lograrse su objetivo, mientras que el del [otro] sí, es irrelevante en relación a cuán condenables son.” (p. 561)

Sin embargo, cabe seguir insistiendo, hay algo moralmente horrendo, que es causar la muerte de alguien, que uno produjo y el otro no, y que es precisamente por lo que el *intento* es también moralmente censurable. O no. También es ciertamente horrendo *tratar de* matar a alguien. Pero de hecho *matarlo*, parece que añade algo más, y negando la relevancia moral de esta diferencia, ¿no se estará trivializando la muerte misma, o el hecho de haberla causado? Como Thomas Nagel (1979) afirma, hay algo *moralmente significativo* en lo que distingue a ambos casos. Si bien sigue teniendo su peso la intuición contraria de que lo que distingue a ambos casos, si realmente está más allá del control de los agentes, no puede *marcar una diferencia* respecto a su responsabilidad moral. Que ambos sean igualmente responsables es

⁴ “El grado de la responsabilidad lo es todo; el alcance no cuenta para nada en cuanto a la evaluación moral de los agentes” (Zimmerman 2002, p. 568). Que el desafortunado deba ser censurado por *más cosas* que el afortunado (lo cual nos muestra fehacientemente que debe ser evaluado negativamente en esta ocasión), no implica que deba ser evaluado negativamente *en mayor medida* que el desafortunado (1987, p. 561). Según este planteamiento, el “alcance” es relevante para otro tipo de juicios morales (los *deónticos* o los *aretáicos*) que son ajenos al merecimiento último del agente.

una consecuencia de aceptar que no podemos ser moralmente responsables de lo que no está bajo nuestro control –de que ninguno de ellos tenía más control sobre las consecuencias de sus actos. En palabras de Michael Zimmerman:

Lo mismo que con la responsabilidad, ocurre con el control [... Uno] tuvo *control de más cosas* que [el otro] (su control tuvo un alcance mayor), pero no tuvo *más control* de lo que sucedió del que tuvo [el otro] (tuvo el mismo grado de control). En tanto que el grado de responsabilidad rastrea el grado de control, [ambos] deben ser declarados igual de moralmente responsables. (Zimmerman 2002, p. 562)⁵

Ciertamente, hay un sentido poco discutible en el que la censurabilidad de ambos es semejante, aunque uno no matara a nadie. Y no se trata solo del hecho de tener la intención, e intentarlo, sino del nivel de elaboración y realización de intención e intento. Su intención firme les llevó a emprender los cursos de acción necesarios para alcanzar su meta: consiguieron el arma, la cargaron y prepararon, investigaron a la víctima, fueron a su encuentro, se le aproximaron, buscaron el momento idóneo, le dispararon, etc. Y, en este sentido, ambos parecen igual de viles.

Podríamos decir, pues, que lo que cuenta es fundamentalmente la intención y el intento (serio) o tentativa. ¿Pero son éstos siempre realmente tan centrales? Comparemos los casos anteriores con un nuevo ejemplo. Pensemos en una madre que cruza la calle por un paso de peatones con su hija pequeña cogida de la mano. Un coche que circula a toda velocidad no respeta la señalización y no frena, atropellando a la niña pero no a la madre. El coche se detiene unos metros más allá. La mujer, que ve a su hija tendida en medio de la calle llena de sangre, es invadida por el odio más furibundo y se aproxima callada al conductor del coche. El joven, queriendo ofrecerle consuelo, al tiempo que expresando sinceramente su pesar, trata de ayudar. Pero ella se alabanza sobre él y trata de clavarle unas tijeras en el estómago. La historia tiene dos finales alternativos, ambos más allá del control de la mujer. En los dos finales la niña se salva sin grandes secuelas. Pero en el primero el joven resulta herido de muerte, mientras que en el segundo el intento se ve frustrado por los reflejos de este, que consigue echarse atrás y bloquear la mano de la mujer. Lo que me interesa es el juicio que merecería la mujer en cada uno de los resultados. La intención es manifiesta sea cual sea el resultado. Pero parece que en el primer final la mujer merecería una condena más severa que en el segundo. La mujer tiene el derecho a reaccionar con rabia y odio ante la conducta del joven, así como de expresar ostensiblemente su reacción. Pero asesinarlo no es una respuesta a la que

⁵ Afirma Zimmerman que, en el sentido pertinente de “control”, un agente puede tener control sobre una cosa aunque suceda por causa de alguna otra cosa que el agente no controla. Reconoce que lo incontrolable es en gran medida ineliminable de nuestras vidas, pero eso no implica que marque una diferencia moral, ni que no podamos tener control alguno.

esté moralmente autorizada. Y si lo intenta pero no lo consigue, las circunstancias restarán importancia a este intento. Pero si lo consigue, si lo mata, las circunstancias no parecen poder mitigar la seriedad del resultado.

Sin duda, se puede replicar que aquello que tiene de condenable la actitud de la mujer no es, de hecho, matar o no al hombre, sino *querer* matarlo –la intención (y el intento) de matarlo–, que es igual en ambos casos y, en principio, no depende de cómo resulten las cosas.⁶ Pero así, la mujer merecería igual condena en ambos casos, cosa que choca con la presumible diferencia de reacciones que nos provoca uno u otro resultado. La diferencia en los resultados presumiblemente cambiará de forma más significativa la consideración moral que el agente merece. Y creo que la diferencia puede apuntar a cómo la realización o no de intención e intento expresan y redefinen su carácter. En el primer final, la mujer se convertirá en una *asesina*, mientras que eso no sucederá en el segundo, en el que la cuestión será más rápidamente liquidada. (El caso parece paralelo al de los asesinos anteriores: si entre el asesino “exitoso” y el asesino frustrado solo hubo la diferencia de la falta de puntería, o la intromisión de un obstáculo en la trayectoria de la bala, ahora entre el primer y el segundo final solo hay la reacción más o menos rápida del joven. Pero la reacción que nos provoca uno y otro caso no es exactamente la misma. En el caso de la mujer la condena depende más del resultado.)

Hay que remarcar que la alternativa que propongo no niega que las intenciones o voluntad del agente jueguen un papel importante en la asignación de responsabilidad moral; sino solo que las intenciones y la voluntad sean el único elemento a considerar, o siempre el más importante. Esto puede resultar más convincente en relación con los casos de negligencia o temeridad que consideraré en la sección siguiente. Del conductor bebido que acaba atropellando a alguien no puede decirse que tuviera la *intención*, ni que *intentara*, matar a nadie –al menos, no en el sentido ordinario de “intención” y de “intentar”, que sí reconocemos en el caso de los asesinos–; por lo que en casos de negligencia, la asignación de responsabilidad no puede depender de la intención e intento del agente, sino de otros factores.

3. Previsibilidad, riesgo y negligencia

Consideremos con más detenimiento el caso de dos conductores ebrios. Dos personas se emborrachan en un bar y deciden volver a casa conduciendo.

⁶ Podría replicarse que la acción es producto de una pérdida momentánea de lucidez: la mujer estaría fuera de sí, no controlaría lo que hace, y por lo tanto no sería moralmente responsable de ninguno de los dos resultados posibles, pues no lo sería ni de su misma acción. Pero, entonces, puedo modificar el caso. Por ejemplo, la mujer maquinó la venganza fríamente mientras su hija estaba en coma en el hospital debatiéndose entre la vida y la muerte. Aquí, no podríamos hablar de pérdida momentánea del control volitivo.

Supongamos que su decisión de conducir tras haber bebido, aún a pesar de los efectos del alcohol, es libre. Con ella ambos cometen una negligencia; esto es, asumen un riesgo moralmente injustificado, en este caso, con respecto a la vida de otras personas.⁷ El defensor del primado de la voluntad (PV) no necesita saber más; aquí termina la historia en lo que a la evaluación moral se refiere. Esta dependerá de la *previsibilidad o probabilidad de que se den ciertas consecuencias*, dada la información relevante accesible al agente en el momento de la acción.⁸ Ambos conductores ebrios son igualmente negligentes, pues ambos someten a los demás usuarios de la vía pública a un riesgo moralmente injustificado. Conducir ebrio es, claramente, una acción en la que los daños potenciales son (en condiciones normales) mayores que los posibles beneficios. Además, el atropello o accidente automovilístico es una consecuencia previsible del hecho de conducir bajo la influencia del alcohol, y esto es algo que normalmente la gente sabe o, en todo caso, debe saber.⁹

Sin embargo, una parte importante de nuestra asignación cotidiana de culpabilidad parece surgir del mismo hecho de haber atropellado a una persona, más que exclusivamente de la mera temeridad del conductor —lo cual no sucede, o no sucede de la misma manera, en el caso de los asesinos. No obstante, el que se ha dado en llamar *argumento epistémico* nos brinda una explicación de este fenómeno que armoniza con el primado de la voluntad: si juzgamos a un conductor de modo diferente al otro no es porque ambos *merecan* un juicio diferente, sino porque su conducta no nos muestra con la misma claridad que lo que merecen sea igual. No siempre está claro cuáles son las intenciones de una persona, ni su compromiso para con un curso de acción determinado. Por ello, el atropello es el mejor indicador que tenemos para tratar a uno de los conductores más duramente. A este respecto, Norvin Richards afirma que si un conductor no causa ningún daño puede ser porque estaba suficientemente alerta o porque su velocidad era segura; o quizá no, y no estaba alerta o su velocidad era excesiva, o ambas cosas. Pero, en cualquier caso, si no ha habido ningún daño real la cuestión permanece incierta. Es solo cuando este conductor atropella a alguien que se torna claro que su atención no era suficiente o que su velocidad era excesiva; en definitiva, que no sopesó correctamente el riesgo de daño a otras personas.¹⁰ Sin embargo, este tipo de argumento hace depender muy

⁷ Se asume un riesgo injustificado cuando el daño potencial de la acción pesa más que sus beneficios potenciales, dadas las probabilidades. Ver Domski (2005), p. 533.

⁸ Ver Rosebury (1995), pp. 518-519; y también Enoch y Marmor (2007).

⁹ Cabe notar que la aceptación de que lo relevante para los casos de negligencia es la previsibilidad o probabilidad de las consecuencias supone ya una cierta renuncia en la defensa del primado de la voluntad; pues, como adelanté, no puede decirse que quien daña por negligencia *quisiera* realmente dañar, por lo menos en un sentido inmediato de querer. Por lo tanto, el principio del primado de la voluntad, contra el que argumento, debe entenderse de un modo más amplio, que no es tampoco incompatible con ciertos compromisos consecuencialistas.

¹⁰ Ver, en particular, Richards (1986), p. 171-2.

estrechamente el merecimiento del agente de las situaciones que podemos llamar (supuestas) “situaciones epistémicamente *transparentes*”, esto es, las situaciones en las que el agente produce de hecho un daño: el conductor ebrio que provoca un accidente, el asesino que consigue acabar con la vida de su víctima, etc. Lo que se afirma es que la producción real de determinado resultado –un daño (o un beneficio, en el caso del elogio)– es un indicador fiable del compromiso del agente con dicho resultado. Contra esto cabe decir lo siguiente. Sin duda, las intenciones del agente no están a menudo claras (incluso para el agente mismo), como puede no estarlo su compromiso con el curso de acción seguido, y normalmente estamos en una situación epistémica mejor para adscribir una temeridad cuando hay un accidente que cuando no. Sin embargo, producir un resultado dañino puede ser tan *accidental* con respecto a las intenciones del agente como lo puede ser intentarlo y no conseguirlo. El atropello no tiene por qué mostrar que el conductor tomara una actitud temeraria. Quizá con una evidencia más clara, como sería el caso si se hubiera podido contar con un registro de velocidad que nos indicara que el coche no sobrepasó la velocidad recomendada, podría hacerse patente que el merecimiento era bien otro.¹¹ No obstante, Richards reconoce que hay ocasiones en las que, se produzca o no finalmente el accidente de tráfico, el sentido común respecto al tiempo de reacción y visión humana y la naturaleza de unas condiciones determinadas es suficiente para concluir que cualquiera que condujese de esta manera, en este lugar, en este tiempo, conduciría temerariamente.

En todo caso, lo fundamental de esta posición es que hace depender el merecimiento moral exclusivamente de la negligencia o temeridad cometida, con independencia de nuestro acceso epistémico. En el caso de los conductores, la acusación moral debería ser (por decirlo en términos legales) solo por conducción temeraria, sin tener en cuenta el consiguiente homicidio por imprudencia. Sin embargo, contra esta posición puede citarse la intuición contraria de que una parte importante de nuestra asignación de merecimiento parece surgir del hecho mismo de haber causado o no un daño –un accidente, un atropello y un homicidio–, más allá de la temeridad misma. Y dando una vuelta de tuerca más a esta intuición, podría decirse incluso que *en algunos casos parece que la misma negligencia solo existe cuando el resultado es dañino*. Sé que esta afirmación producirá la protesta inmediata de muchos. Pero permítaseme tratar de defenderla con ayuda del siguiente ejemplo.

Una mujer está bañando a su hijo, cuando alguien llama a la puerta. Presumiblemente es su anciano padre, al que estaba esperando. Dado que está sola en casa, decide dejar al niño por un momento para ir a abrir la puerta. Pero al abrir la puerta se encuentra con un vecino que quiere informarla de una avería en la puerta del garaje. La mujer trata de deshacerse educadamente del vecino, pero este insiste en repetir los pormenores de la avería y la entretiene más de la cuenta.

¹¹ Para otra réplica de este tipo, véase Adler (1987), pp. 247-8.

Finalmente, logra deshacerse del vecino, cierra la puerta y vuelve rápidamente al baño. Cuando llega, dos resultados son posibles. En el primero, el niño sigue chapoteando alegremente en el agua. En el segundo –el trágico–, el bebé se ha hundido bajo el agua. Podría decirse que en un caso como este la negligencia o temeridad solo *existe* –o, mejor, solo es significativa– si el resultado es malo. En el escenario en el que la madre vuelve al baño y su niño sigue chapoteando en el agua, no parece razonable hablar de ningún tipo de negligencia importante por su parte. Es solo en el caso en el que el niño sufre daños cuando se puede hablar de una clara negligencia. La explicación de que la negligencia o temeridad solo *exista* o sea significativa cuando el resultado es malo reside en el hecho de que el resultado dañino tiene el efecto de elevar el *estándar de exigencia*, de anticipación o precaución debida, para esa situación determinada. Y la explicación de esta diferencia podría ser la siguiente: comúnmente todos tomamos riesgos, riesgos morales incluso, que en el día a día pueden resultar inevitables, pero supondría una concepción demasiado rigurosa de la moralidad y la vida en general, y con toda seguridad impracticable, igualar la censura moral que la madre merece en ambas situaciones, con independencia del resultado de hecho. El estándar de exigencia no puede ser cotidianamente tan alto. Por ello, no podemos juzgar a la madre con la misma dureza en ambos casos, pero tampoco podemos tratarla con la misma indulgencia.

Veamos aún otro caso que apunta en la misma dirección. Se trata de un caso de falta de atención o distracción. Un joven acaba el turno de noche en el hospital y conduce hacia su casa. Mientras circula por una gran avenida, entre decenas de coches, enciende la radio y trata de hallar una cadena en la que den las noticias. Le cuesta encontrar una, centra su atención en ello, mira durante unos segundos la pantalla del dial, y entonces oye un golpe contra su coche. Cerca de donde frena, hay una moto en el suelo y junto a ella un motorista que parece malherido. En realidad, ha fallecido.¹² No puede decirse que cometiera ninguna gran temeridad, pero se descuidó. Este descuido podría haber sido de lo más intrascendente, pero no lo ha sido: su coche se desvió un poco de su trayectoria y embistió al motorista que –¡vaya mala suerte!– circulaba a su lado.

No cabe duda de que la misma idea de riesgo (moral) es muy resbaladiza. Imaginemos que alguien acude a una fiesta y se emborracha. Él sabe que cuando está bajo la influencia del alcohol se desenfrena y puede realizar trastadas muy diversas. ¿Cuál es el riesgo al que está sometiendo a los demás asistentes a la fiesta cuando empieza a ingerir alcohol? Pongamos que se acerca a alguien y le echa encima su cubata. O imaginemos que, para gastar una broma, coge a su mejor amigo y trata de tirarlo a la piscina desde el balcón de la casa. Y, finalmente, tras un forcejeo, consigue arrojar a su amigo por el balcón y, por supuesto, no llega hasta la piscina. Supuestamente, desde el punto de vista del primado de la voluntad, pare-

¹² Tomo el ejemplo de Corbí (2003), p. 38.

ce que nuestro protagonista debe ser censurado igualmente en el caso de que cualquiera de estos resultados se dé, pues lo que importa es exclusivamente el riesgo a que injustificadamente somete a los demás, y este es el mismo en ambas situaciones. No hay duda (y esta es mi posición) de que, en tanto que responsable de emborracharse (y también de tratar de lanzar a otra persona a la piscina desde el balcón), debe cargar con las consecuencias, inmediatas o de sus acciones subsiguientes, pero no solo en virtud del riesgo general o abstracto al que somete a los demás, sino también en virtud de las consecuencias particulares que *de hecho* se producen. ¿Cómo justificar si no la diferencia en culpabilidad de una situación (verter el cubata sobre alguien) a otra (arrojar a alguien por el balcón), si el *locus* de la responsabilidad es meramente la negligencia de excederse en la ingesta de alcohol en una fiesta? ¿O es que estamos dispuestos a afirmar que, en ambos casos, la culpabilidad es la misma? Por otro lado, para que esta posición pudiera concretarse se debería poder definir con exactitud y *ex ante* la idea de riesgo (de qué supone someter a otros a un riesgo injustificado), sin reinterpretarse en virtud de alguna consecuencia determinada de la acción evaluada –que es lo que parece suceder en la interpretación de los casos particulares considerados. Si bien con esto no quiero negar, por supuesto, que haya casos en los que el riesgo es más o menos claro y determinable con independencia de los resultados o acciones subsiguientes.

Pero, ¿qué hay de las decisiones en condiciones de incertidumbre? Veamos un ejemplo. Los líderes de un grupo opositor clandestino deben decidir cuándo atacar al gobierno del cruel régimen dictatorial que subyuga el país. La decisión de lanzar finalmente la insurrección no puede basarse en el conocimiento cierto de cómo acabará todo, pues no existe este conocimiento. Inevitablemente, estamos ante una decisión que implica un gran riesgo, y los factores que pueden influir en el resultado son muy variados. Si la insurrección tiene éxito, se obtendrá la libertad; pero, por el contrario, también puede suceder que el intento quede frustrado y los dirigentes del régimen tiránico respondan con una represión sangrienta que provoque un coste muy grande en vidas y sufrimiento. Así, en muchos casos, la elección se torna muy difícil, a causa del hecho de que el resultado no puede ser previsto con certeza. O imaginemos aún el caso de un experto en negociación que, arriesgando la vida de los rehenes, se niega a satisfacer las demandas del secuestrador, *convencido* de que con ello logrará que este se derrumbe. Puede que crea que tiene buenas razones para hacer esto, pero el riesgo es enorme. Si la estrategia sale bien y el secuestrador se acaba derrumbando y se entrega, parece que el negociador tendrá poco o nada que reprocharse; pero si resulta que el secuestrador era psicológicamente más resistente de lo que el negociador intuía y responde matando a los rehenes, la actitud del negociador hacia su propio comportamiento cambiará, o debería cambiar.

No hay duda de que este tercer tipo de casos está muy estrechamente conectado con el anterior. Tanto en las situaciones de decisión en condiciones de incertidumbre como en los casos de negligencia, el defensor del primado de la voluntad

circunscribe la asignación de responsabilidad moral a la decisión misma de actuar, que se identifica con la perspectiva *ex ante*. Esta perspectiva toma exclusivamente en consideración la previsibilidad o probabilidad de las consecuencias en virtud de la información relevante conscientemente accesible al sujeto en el momento de deliberar y decidir. Por el contrario, mis últimas objeciones apuntan a la significación que la perspectiva retrospectiva también tiene en la asignación de responsabilidad moral. Sin duda, esta es la conocida conclusión de Williams (1981); si bien mi argumento se basa, en concreto, en la idea de la variabilidad del estándar de exigencia según la gravedad moral del resultado.

4. Lo voluntario frente a lo inadvertido

Por otro lado, en línea con la defensa del primado de la voluntad (PV), Harry Frankfurt afirma lo siguiente en relación a lo que una persona hace o produce *inadvertidamente*:

La conducta que no producimos deliberadamente es conducta *no del todo voluntaria*: nosotros no *pretendemos* realizarla del todo. Es natural pensar que sería *injusto considerarnos moralmente imputables o responsables* –esto es, considerarnos objetos legítimos de *estima moral* y *oprobio moral*, o de *elogio moral* y *censura moral*– por aquella conducta en la que nos vimos involucrados, pero que no *emprendimos voluntariamente*. (Frankfurt 2009, p. 12; traducción levemente modificada.)

Sin duda esta distinción entre lo voluntario y lo inadvertido, o entre la acción deliberada y la conducta que no intentamos realizar deliberadamente, es de sentido común, y es de hecho necesaria para cualquier caracterización satisfactoria de la acción humana. Ciertamente, toda concepción de la responsabilidad moral trata de establecer una distinción clara entre aquello por lo que podemos ser moralmente responsables y aquello por lo que no. Sin embargo, el desacuerdo empieza cuando tenemos que definir exactamente esta distinción y, más aún, cuando tratamos de aplicarla a casos concretos. Frankfurt compara los efectos no directamente deliberados de nuestras acciones con los movimientos espasmódicos producidos por un ataque epiléptico incontrolable. El problema es que parece claro que la diferencia, por ejemplo, entre tratar de rescatar a alguien de las llamas de un incendio y el resultado de terminar arrojándolo de un duodécimo piso en el intento no es análoga a la diferencia entre controlar con normalidad nuestras acciones y ser víctima de un ataque epiléptico.¹³

¹³ El ejemplo del bombero, procedente de Nagel (1979), es discutido por Frankfurt (2007), *passim*. Frankfurt también contrasta, de manera igualmente desafortunada, el producir un daño intencional-

Obviamente, el ser víctima de un ataque epiléptico es algo por lo que uno no es moralmente responsable, incluso si en el transcurso de este ataque alguien resulta dañado, si es el caso que uno no podía reprimirlo de ninguna manera. Pero sí que estaba en manos del bombero (o lo estuvo anteriormente) ser más hábil en el rescate. En general, la cuestión es que el control que ejercemos sobre nuestras acciones es muy variable; es de muy diversos grados. El ataque epiléptico es un caso extremo en el que el control del agente parece estar totalmente ausente y, por lo tanto, es tramposo o engañoso igualarlo a otro tipo de acción en la que, si bien el control del agente es reducido, o indirecto, este no parece estar totalmente ausente. Por otro lado, defender que también somos moralmente responsables por estas últimas acciones no significa que se deban desestimar, para la determinación del juicio de responsabilidad moral concreto, las diferencias graduales en el control que ejercemos sobre ellas. Por supuesto, no somos responsables en la misma medida –esto es, la gravedad de nuestra responsabilidad es distinta– en relación a, por ejemplo, matar intencionalmente a alguien, intentar matarlo, hacerlo a causa de una negligencia, o por accidente.

5. Intención, carácter y acción

Hasta ahora, mi rechazo del primado de la voluntad (PV) se ha limitado a tratar de mostrar que factores que van más allá de los motivos e intenciones deliberados del agente pueden establecer (porque en ocasiones lo hacen) una diferencia en el juicio de responsabilidad moral que una persona merece. En esta última sección, sin embargo, me centraré en la cuestión de la relación entre intención y acción. Me gustaría mostrar, aunque solo sea de un modo parcialmente esquemático, (i) que las circunstancias en las que nos hallamos suelen marcar una distinción con respecto de nuestras intenciones o disposiciones previas, en tanto que nos abren ciertos cursos de acción determinados; y (ii) que el hecho de llevar a cabo cierta acción puede contribuir significativamente a la fijación misma de la intención por la que el agente realiza la acción o, en todo caso, suele conllevar un refuerzo importante del compromiso del agente con dicha intención.

Tomemos el ejemplo de una persona que llega a cometer una infidelidad aún a pesar de que anteriormente al encuentro con la persona con la que cometió dicha infidelidad no pretendía ser infiel. Supongamos que esta descripción es correcta y que no se da ningún tipo de autoengaño en el agente respecto a sus planes o disposiciones anteriores. En este caso, puede parecer que uno es responsable de una infi-

mente con ser portador de una enfermedad tóxica contagiosa que daña a los que nos rodean sin que podamos hacer nada para impedirlo (secc. 9), como análoga a la distinción entre lo voluntario y lo inadvertido.

delidad que no habría cometido si no hubiera estado en las circunstancias propiciatorias para ello. Si bien, podemos tender a eximir a esta persona si consideramos que fueron las circunstancias las que la llevaron a ser infiel. Imaginemos, por ejemplo, que la infidelidad se produjo después de una cena de negocios en la que bebieron en abundancia. Ciertamente, esto podría excusar al agente en algunos casos, pero aún así no parece lo más habitual. En muchos casos tendemos a considerar a la persona culpable por su infidelidad, aun siendo cierto que de no haberse vuelto envuelta en esa cena muy probablemente no habría sido infiel, ni podemos atribuirle la intención o plan previo de ser infiel. Es cierto que alguien puede replicar que la responsabilidad se halla más bien en la *negligencia* de no haber evitado una situación que podía derivar en infidelidad; por ejemplo, la negligencia de haber aceptado ir a cenar con cierta persona, cuando podía sospecharse que ese resultado –la infidelidad– era posible o probable (véase Doris 2002, p. 144). Sin embargo, esto no es del todo claro: ir a la cena no es en sí mismo ni negligente, ni directamente incorrecto. Más bien, ir a la cena parece ser condenable solo si resulta en infidelidad; lo que apunta a la conclusión de que lo condenable es, fundamentalmente, cometer la infidelidad misma. Normalmente diríamos que si el agente es culpable lo es principalmente por el hecho mismo de cometer la infidelidad, no por no haber evitado la ocasión propiciatoria de esa infidelidad. Esto último es una mera cuestión estratégica que queda para el agente, sin repercusión alguna en el juicio.

Otro tanto sucede, sin duda, en los casos en los que las circunstancias permiten al agente obtener cierto elogio, en contraste con sus intenciones o disposiciones anteriores. Piénsese en el caso de un anciano que odia sinceramente a los inmigrantes y sostiene que todos ellos deberían ser expulsados del país. No será muy difícil ponerle una cara familiar a esta persona, pues todos conocemos a personas que mantienen opiniones más o menos extremas respecto a ciertos temas sociales y políticos; opiniones que defienden sinceramente, e incluso están convencidos de que si estuvieran en los puestos de decisión adecuados no dudarían en actuar en consecuencia. Escojamos a un hombre que rechaza tan frontalmente a los inmigrantes que no soporta ni cruzarse con ellos; pero que un día tiene que hacer cola en una oficina de desempleo y se ve rodeado de inmigrantes. Al entrar en contacto directo con alguno de ellos, se muestra, no obstante, cordial e incluso se preocupa por uno al que le ha dado un pisotón. Tras este hecho, el juicio que merece el anciano parece ser menos severo, pues ya no solo *odia* a los inmigrantes, en abstracto, sino que también vemos que cuando se ha encontrado con algunos los ha tratado correctamente. Así, el hecho de haberse encontrado con algunos inmigrantes en la cola del paro parece marcar una diferencia en el juicio que el agente merece; que es ahora menos adverso. Es el hecho mismo de verse envuelto en ciertas circunstancias y responder de determinada manera lo que *crea* un mérito que va más allá de sus meras disposiciones (o disposiciones anteriores).

Un caso paralelo, o que al menos apunta en la misma dirección, es el de la conducta del personaje de Mark Twain, Huckleberry Finn, cuando ayuda a escapar al esclavo Jim, contra sus mismas *convicciones morales*. Es el hecho de afrontar unas circunstancias determinadas –la disyuntiva de denunciarlo o permitir que se escape– lo que hace que descubramos el mérito moral de Huck, que va más allá de sus mismas creencias sobre sus propios valores morales acerca de la esclavitud –que son los imperantes en su sociedad y que es incapaz de cuestionar o revisar.¹⁴ En general, el mérito de resistirse a una tentación, o el demérito de sucumbir a ella, o de aprovechar o no una ocasión propicia, solo es posible si la tentación o la ocasión propicia se nos presentan, y cuando se nos presenta.

Ciertamente, el defensor del primado de la voluntad tiene que negar que la mera aparición de la circunstancia u ocasión pueda realmente marcar una distinción moral; a lo sumo puede marcar una *distinción epistémica*. Tras ver al anciano conernido por el daño que el inmigrante pueda sufrir, *descubrimos* (como él mismo hace) que su actitud no era exactamente la que parecía ser. Algo parecido se podría decir de la persona infiel. Sin embargo, esta es una manera demasiado superficial de entender la conducta humana. Parece cierto, por un lado, que respecto a muchos rasgos, actitudes, creencias, deseos, etc., cabe esperar a su manifestación conductual para poder juzgar correctamente a su poseedor. Y esto es cierto incluso en el caso de uno mismo –lo que, sin duda, tiene que ver con los límites del autoconocimiento. Como muchos experimentos psicológicos muestran, a menudo ignoramos los motivos reales de nuestras acciones, y de hecho sucede en muchas ocasiones que el que consideramos *el* motivo por el que llevamos a cabo cierta acción no es más que el producto de una mera *confabulación* o *racionalización*.¹⁵ En general, es sencillamente presuntuoso suponer que podemos saberlo todo acerca de los motivos e intenciones de una persona, incluido uno mismo. Nuestros móviles no son nunca totalmente transparentes, ni a los demás, ni a nosotros mismos.

Pero nuestro adversario puede mantenerse firme replicando que seguimos ante un problema meramente epistémico: ciertamente, nuestra comprensión del carácter de una persona puede ser difícil y estar oscurecido por los hechos, pero eso es independiente de cuál sea realmente la maldad o bondad misma de su carácter. Así, podría replicarse que lo que muestra la conducta del anciano anterior es que aquel hombre no odiaba realmente a los inmigrantes, sino que simplemente estaba equivocado acerca de su verdadera actitud para con ellos, que no depende solo de su reflexión y convencimiento, sino también de sus reacciones espontáneas y/o afectivas.

¹⁴ Véase Bennett (1974) para una interesante discusión de este caso, que ha dado origen a cierta bibliografía. Arpaly y Schroeder (1999) llaman a esto “acrasia inversa”; véase también Arpaly (2003), cap. 3.

¹⁵ Esto es especialmente chocante en los estudios sobre los efectos en la conducta de la manipulación del estado de ánimo en que uno se encuentra, o también en las diferencias en nuestras compras debido a la colocación de los productos en un supermercado, por ejemplo.

No obstante, con esto el defensor del primado de la voluntad sigue haciendo concesiones, pues las reacciones espontáneas y/o afectivas de alguien dependen de las circunstancias y, en general, superan los límites de la voluntad, en sentido estricto. Además, a mi entender, la cuestión es más bien que centramos exclusivamente en la intención o volición previa del agente, aquella que podríamos incluso registrar si tuviéramos acceso directo a su mente en el momento de decidirse a actuar, puede llevarnos a error. Pues es nuestra misma conducta en muchas situaciones la que nos hace *aclararnos* acerca de nuestras intenciones o motivos, lo cual solo es posible a partir de la prueba que supone estar en unas determinadas circunstancias y de acabar actuando de una u otra manera. No se trata de un mero descubrimiento o incluso reinterpretación de nuestras intenciones a la luz de la acción efectiva, sino que la acción misma es parte constitutiva de la intención o rasgos internos —esto es, contribuye en su fijación. No es que antes de aquella acción el hombre interpretase erróneamente su actitud y ahora haya llegado a la interpretación correcta. (Esto puede suceder, pero no es necesario que suceda.) Sino que muchos estados mentales o rasgos del agente están sencillamente indeterminados hasta que ciertas conductas efectivas contribuyen a su fijación.

Alguien que nunca se ha encontrado en cierta situación, como por ejemplo una situación real de guerra, que reclama la manifestación de cierto rasgo, como el valor (propio de la guerra), no solo no conoce cuál sería su valor o falta de él en esas circunstancias, sino que parece que en realidad su carácter ni posee este rasgo ni el contrario, sino que es más bien *indeterminado* respecto a él. De hecho, hablar de carácter presupone hábitos, *maneras características* en que una persona cree, siente y actúa, disposiciones *normales* a actuar y sentir; que respecto de tal rasgo no existen.

Esta tesis puede subsumirse en la idea más general de que el carácter del agente no es, en general, aislable de su conducta efectiva, sino que está parcialmente constituido por esta. ¿Pero, es esto así? ¿No puede ser el caso que una persona sea, por ejemplo, cobarde, sin que haya manifestado externamente nunca este rasgo? ¿O incluso aunque nunca haya experimentado internamente una tendencia que cabría calificar de actitud cobarde? Esta última posibilidad parece demasiado radical. Resulta absurdo mantener que uno puede tener un rasgo de carácter aun sin que nunca lo manifieste, ni interna ni externamente.¹⁶ Por otro lado, la posibilidad de un contraste radical entre la inclinación interna y la conducta externa es, cuanto menos, altamente inusual en el mundo real; si bien podría darse el caso de que alguien consiguiera refrenar siempre su inclinación a la envidia, por ejemplo. Alguien que es envidioso, pero que en cierta ocasión evita comportarse envidiosamente, no deja

¹⁶ No pretendo negar que alguien pueda ser cobarde, o tener cualquier otra disposición conductual, sin saberlo; pero entonces habrá algo en su conducta que indique la existencia de esta disposición, aunque él no se percate.

por ello de ser envidioso y, a este respecto, se le puede censurar por ello –en particular, por haber adquirido, o no acabar con, esa disposición–; aunque, por supuesto, si alguien logra resistir su tendencia a la envidia y no actuar envidiosamente merece un elogio que no merece el que simplemente se deja llevar por ella. Esto es, normalmente valoramos que alguien consiga resistir una inclinación negativa –como también valoramos que alguien realice de hecho, convierta en acto, cierta inclinación positiva. Es por ello que fundar la atribución de responsabilidad moral exclusivamente en los estados mentales es pasar por alto que existe una diferencia muy significativa entre su mera posesión y su plasmación en la conducta. La idea puede generalizarse así: convertir en conducta cualquier tipo de tendencia, intención o volición es significativamente distinto de solo poseerlas internamente, sean buenas o malas.

6. Conclusiones

En este artículo he presentado diferentes tipos de casos que vienen a contradecir el principio que bauticé como *primado de la voluntad* (PV). En la sección primera me ocupé del contraste entre la intención del agente y las consecuencias de su acción; mientras que en la segunda me centré en los casos de negligencia y de decisión en condiciones de incertidumbre. En la sección tercera consideramos la distinción entre actos realizados deliberadamente y actos realizados inadvertidamente y rechacé que Frankfurt muestre que esta distinción sea fundamental con respecto a los límites de nuestra responsabilidad moral. Finalmente, en la sección cuarta traté de poner en cuestión la distinción nítida entre la intención y la misma acción que la realiza, asumida por el defensor del primado de la voluntad.

Mi meta no ha sido negar la importancia de la voluntad (intención, motivo) en la determinación del juicio de responsabilidad moral que el agente merece, sino la tesis de que es lo *único* que debe tenerse en cuenta. Es importante notar que mi argumento solo aspira al reconocimiento de que los distintos elementos que van más allá de la mera voluntad del agente (o algunos de ellos) introducen matices, constituyen agravantes o atenuantes, o incluso contribuyen a conformar las intenciones mismas del agente; de modo que influyen (en diferente medida) en el juicio de responsabilidad moral merecido.¹⁷

¹⁷ El presente ataque al que he llamado *primado de la voluntad*, que es una forma de vindicación de la suerte moral resultante, será más convincente para aquellos que, de alguna manera, ya acepten que es imposible impedir por completo que la suerte se inmiscuya en el juicio moral –como he tratado de mostrar en Rosell (2012). Con todo, aunque se acepte que la suerte moral, en general, es inevitable, se puede, al mismo tiempo, rechazar la suerte resultante, en particular; pues puede haber buenas razones para no aceptar un determinado tipo, que no se apliquen en general. Es por esto que la defensa de la

Referencias bibliográficas

- ADLER, J. (1987): “Luckless Desert is Different Desert”, *Mind*, 96, pp. 247-249.
- ARPALY, N. (2003): *Unprincipled Virtue. An Inquiry into Moral Agency*, Oxford, Oxford University Press.
- ARPALY, N. y T. SCHROEDER (1999): “Praise, Blame and the Whole Self”, *Philosophical Studies*, 93, pp. 161-188.
- BENNETT, J. (1974): “The Conscience of Huckleberry Finn”, *Philosophy*, 49, pp. 123-134.
- CORBÍ, J. (2004): *Un lugar para la moral*, Madrid, Antonio Machado Libros.
- DOMSKY, D. (2005): “Tossing the Rotten Thing Out: Eliminating Bad Reasons not to Solve the Problem of Moral Luck”, *Philosophy*, 80, pp. 531-541.
- DORIS, J. (2002): *Lack of Character*, Nueva York, Oxford University Press.
- ENOCH, D. y A. MARMOR (2007): “The Case Against Moral Luck”, *Philosophy and Law*, 26, pp. 405-436.
- FRANKFURT, H. (2010): “Inadvertencia y responsabilidad moral”, traducción de Fabio Fang, *Ideas y Valores*, 141, pp. 11-24.
- KANT, I. (1785/1986): *Fundamentación de la Metafísica de las Costumbres*, trad. de Manuel García Morente, Madrid, Espasa-Calpe, 6ª ed.
- NAGEL, T. (1979): “Moral Luck”, en T. Nagel, *Mortal Questions*, Cambridge, Cambridge University Press, pp. 24-38. Se usa la traducción de Williams y Nagel (2013).
- RICHARDS, N. (1986): “Luck and Desert”, *Mind*, 65, pp. 198-209. Se cita la reimpresión en D. Statman, comp. (1993).
- ROSEBURY, B. (1995): “Moral Responsibility and Moral Luck”, *Philosophical Review*, 104, pp. 499-524.
- ROSELL, S. (2006): “Nagel y Williams acerca de la suerte moral”, *Revista de Filosofía*, 31(1), pp. 143-165.
- ROSELL, S. (2012): “El argumento contra la suerte moral: articulación y respuesta”, *Crítica. Revista Hispanoamericana de Filosofía*, 44 (130), pp. 3-33.
- RUSSELL, P. (1999): “Smith on Moral Sentiment and Moral Luck”, *History of Philosophy Quarterly*, 16, pp. 37-58.
- SMITH, A. (1790/1976): *The Theory of Moral Sentiments*, edición de D. D. Raphael y A. L. Macfie, Oxford, Clarendon Press. [Trad. castellana: *Teoría de los sentimientos morales*, trad. de Carlos Rodríguez Braun, Madrid, Alianza, 2004.]

suerte resultante, en particular, requiere de argumentos independientes a los de la defensa de la suerte moral en general; del tipo de los que he tratado de ofrecer aquí.

Este trabajo se encuadra en el proyecto de investigación “Alternativas, creencia y acción” (FFI2009-09686), financiado por el Ministerio de Ciencia y Tecnología del Gobierno de España. Y su redacción fue también posible gracias a la ayuda económica del Programa VALi+d (fase postdoctoral) de la Conselleria d’Educació de la Generalitat Valenciana.

- STATMAN, D., comp. (1993), *Moral Luck*, Albany, State University of New York Press.
- WILLIAMS, B. (1981): “Moral Luck”, en B. Williams, *Moral Luck: Philosophical Papers 1973-1980*, Cambridge, Cambridge University Press, pp. 20-39. Se usa la versión castellana de Williams y Nagel (2013).
- WILLIAMS, B. y T. Nagel (2013), *La suerte moral*, introducción y traducción de S. Rosell, Oviedo, Krk Ediciones.
- ZIMMERMAN, M. (1987): “Luck and Moral Responsibility”, *Ethics*, 97, pp. 374-386. Se cita la reimpresión en D. Statman, comp. (1993).
- ZIMMERMAN, M. (2002) “Taking Luck Seriously”, *Journal of Philosophy*, 99, pp. 553-576.

Sergi Rosell
Phronesis – Analytic Philosophy Group
Departamento de Metafísica y Teoría del Conocimiento
Facultad de Filosofía y CCEE
Universidad de Valencia
sergi.rosell@uv.es