

RESPONSABILIDAD, AUTORÍA Y VOLUNTAD: LA PROPUESTA DE MORAN

MIRANDA DEL CORRAL

CONICET (Argentina)

Dirección: Instituto de Filosofía 'Alejandro Korn', Puán 480, 4º Piso, 1406

Ciudad de Buenos Aires, Argentina

miranda.delcorral@filo.uba.ar

miranda.delcorral@gmail.com

Received:18.04.2014; Revised:09.01.2015; Accepted:22.01.2015

Resumen: Este artículo plantea tres objeciones a la propuesta de Richard Moran acerca de la relación entre razones y responsabilidad. En primer lugar, mostraré que la relación entre razones y responsabilidad que presenta Moran es problemática, por dos razones: primero, las creencias alienadas del agente pueden ser parte del razonamiento; y segundo, las llamadas *razones de estado* no cumplen el Principio de Transparencia, pero no eximen de responsabilidad. En segundo lugar, argumentaré que la propuesta atribucionista de Moran postula una relación demasiado exigente entre estados mentales y razonamiento. Tercero, señalaré la dificultad de la propuesta de Moran para recoger la intuición de que tanto la autoría como la responsabilidad se presentan en grados.

Palabras clave: Responsabilidad, Razones normativas, Autoría, Richard Moran.

Abstract: This article presents three objections to Richard Moran's proposal concerning the relation between reasons and responsibility. First, I will show that the relation between reasons and responsibility suggested by Moran is problematic, for two reasons. On the one hand, alienated beliefs can play a role in the reasoning process. On the other hand, state-given reasons do not meet the Transparency Condition, but they do not exempt the agent either. Second, I will argue that the relation between attitudes and reasons that Moran postulates is too demanding. Third, I will point out that Moran's account cannot accommodate the intuition that both responsibility and authorship come in degrees.

Keywords: Responsibility, Normative reasons, Authorship, Richard Moran.

Introducción

La inmediatez con la que se nos presentan nuestros estados mentales sustenta la tesis de que tenemos autoridad epistémica sobre éstos. Esta autoridad es específica de la primera persona: tenemos autoridad para zanjar la cuestión acerca de qué deseamos o creemos, mientras que debemos inferir qué desean y creen los demás a partir de lo que hacen y dicen. Las teorías del autoconocimiento tratan, entre otras cosas, de dar cuenta de esta autoridad. Richard Moran (2001; véase también 1997, 2002, 2004), analiza la autoridad epistémica en términos de autoridad constitutiva, es decir, de autoría. Según su propuesta, un agente es autor cuando sus actitudes responden a sus razones normativas.

Para Moran¹, autoría y responsabilidad se encuentran estrechamente relacionados. Un agente es *responsable* de sus acciones y de sus actitudes mentales en tanto que es *autor* de ambos. Mientras que el concepto de autoría resulta intuitivamente claro cuando se aplica a las acciones, es necesario elucidar qué significa que un agente sea autor de sus actitudes. En la teoría desarrollada por Moran, la autoría sobre una actitud consiste en una adecuada respuesta a las propias razones para mantener esa actitud. Un agente es *autor* de sus actitudes cuando éstas son susceptibles de responder a sus razones: el agente puede entonces responder de sus actitudes apelando a sus razones. Es decir, está en posición de emprender la tarea de *justificarlas* ante sí mismo y ante otros. La autoría, por lo tanto, implica responsabilidad en un sentido deliberativo. Moran se sitúa así en la línea de las teorías atribucionistas de la responsabilidad moral, que defienden que los agentes son responsables de aquellas acciones y actitudes mentales de los que son

¹ En este artículo, me centraré en el argumento que Moran presenta en *Authority and Estrangement* (2001), aunque también comentaré trabajos de Moran posteriores a esta obra.

autores, en tanto que éstas expresan sus razones².

La finalidad de este artículo es evaluar críticamente la postura de Moran acerca de la relación entre responsabilidad y razones. En la primera sección, expondré los elementos principales de la teoría de Moran, centrándome en la relación entre autoconocimiento y autoría. En lo restante del artículo, señalaré tres objeciones a esta propuesta. Primero (sección 2), mostraré que la relación entre razones y responsabilidad no es tan directa como Moran postula, ya que, por una parte, las creencias alienadas del agente pueden ser parte del razonamiento; y por otra parte, las llamadas *razones de estado* no cumplen el Principio de Transparencia, pero tampoco eximen de responsabilidad. En segundo lugar (sección 3), la propuesta atribucionista de Moran postula una relación demasiado exigente entre estados mentales y razonamiento, ya que para que un agente pueda ser responsable de sus actitudes proposicionales, dice Moran, éstas han de ser la conclusión posible del razonamiento. En contra de esta tesis, argumentaré que esta relación no es ni necesaria ni suficiente para la autoría. Por último, en la cuarta sección, señalaré la dificultad de la propuesta de Moran para recoger la intuición de que tanto la autoría como la responsabilidad se presentan en grados.

² Diversas teorías atribucionistas difieren acerca de en qué consiste exactamente ser autor de las propias acciones y estados mentales. Dentro de estas teorías podemos distinguir entre las que argumentan que ser autor conlleva expresar los propios valores en las acciones, desvelar el “verdadero yo”, y las teoría que aducen que ser autor consiste en responder de determinada manera a las razones que justifican las propias acciones o estados mentales (véase Vargas, 2009 para un análisis comparativo entre estas dos teorías). Como veremos, Moran defiende una teoría atribucionista basada en la capacidad del agente de hacer que sus estados mentales y sus acciones respondan a sus razones; se sitúa por lo tanto en la misma línea que Pettit y Smith (1996), Scanlon (1998), O'Brien (2007) y Hieronymi (2008), entre otros.

1. Autoconocimiento, autoría y responsabilidad

Si bien es cierto que todos nuestros estados mentales nos pertenecen, en el sentido de que todas mis creencias, deseos, miedos y esperanzas son necesariamente míos, no nos relacionamos con todos ellos de la misma manera. Podemos tener deseos enajenados, desarrollar una fobia, o sorprendernos a nosotros mismos deseando algo que nos repugna desear. Estos son estados mentales de los que no somos autores, sino meros espectadores. Podemos conocer estos estados alienados, argumenta Moran, pero desde una perspectiva teórica, de tercera persona. Sin embargo, sólo adoptando una actitud de primera persona podemos *declarar*³ tener estos estados mentales, y ofrecer las razones que los justifican.

Esta autoridad de primera persona es, por una parte, epistémica: los sujetos conocen sus propios estados mentales de una manera inmediata, no inferencial ni mediante evidencia comportamental. Por otra parte, la actitud de primera persona expresa una autoridad práctica: en tanto que agentes, podemos deliberar, decidir y zanjar la cuestión sobre qué creer, qué desear o qué hacer. Conocemos nuestros estados mentales porque los constituimos mediante procesos deliberativos. Y somos autores de nuestras actitudes mentales cuando nuestras éstas responden a nuestras razones, es decir, cuando cumplen el Principio de Transparencia:

One is an agent with respect to one's attitudes insofar as one orients oneself toward the question of one's beliefs by reflecting on what's true, or orients oneself toward the question of one's desires by reflecting on what's worthwhile or diverting or satisfying. (Moran, 2001, p. 64)

³ En inglés, *to avow*; se refiere al acto de comunicar estados mentales, declarándolos como propios.

Así, es nuestra condición de agentes la que nos permite mirar al mundo y buscar las razones a favor de *p*, cuando queremos saber qué creemos al respecto de *p*. Y al mismo tiempo, en tanto que agentes, desempeñamos un papel activo en la formación de estados mentales: decidimos que algo es el caso, o que una acción es deseable. Sin esta capacidad de decisión, afirma Moran, el Principio de Transparencia perdería su sentido:

[O]nly if I can see my own belief as somehow up to me will it make sense for me to answer a question as to what I believe about something by reflecting exclusively on that very thing, the object of my belief. (Moran, 2001, pp. 66–67)

Un agente responsable considera sus creencias una expresión de su criterio de evaluación y ponderación de razones prácticas y teóricas. Sus acciones y estados mentales son sus obras, y no un mero movimiento corporal, o una sensación con respecto a la cual el sujeto tiene un papel pasivo:

To believe X, or to intend X, is plainly to assume certain responsibilities with respect to the reasons in favor of taking X to be true, or the reasons in favor of doing X. Believing X, for instance, involves taking on the responsibility for responding appropriately to evidence against it, and to drawing conclusions from it in conjunction with one's other beliefs. To believe X is to 'bind oneself' with respect to reasons for believing the truth of X, just as in intending X one binds oneself with respect to the future occurrence of X. [...] Such binding of oneself, as in agreement, or its explicit denial in disagreement, is the sort of thing that can reasonably be asked of us, and is the taking on of a kind of responsibility for oneself. (Moran, 2003, p. 416)

La responsabilidad del agente consiste en hacer que sus actitudes respondan adecuadamente a las razones que justifican estas actitudes. De él depende, en un sentido relevante, el mantener, revisar o abandonar una creencia, una intención, o un deseo. Ahora bien, ¿en

qué sentido relevante las creencias dependen del agente?

Algunas actividades mentales pueden ser realizadas de manera voluntaria: por ejemplo, puedo recordar a mis padres, imaginar un oso blanco, o suponer que fuese un cerebro en una cubeta. Otras, sin embargo, tales como juzgar que *p* es el caso, decidir estudiar francés, temer el advenimiento de la III Guerra Mundial, o desear ir a la playa, no parecen depender de la voluntad en el mismo sentido que las primeras. La razón por la que no podemos creer o desear a voluntad, pero sí mover nuestra mano o imaginar un oso blanco, es porque la creencia y el deseo son actitudes que Moran denomina, siguiendo a Scanlon (1998, p. 18 y ss.), “sensibles al juicio”. Estas actitudes son susceptibles de estar justificadas por razones normativas. Dar las razones por las que creemos que *p* consistiría en mostrar los hechos que justifican la verdad de la proposición *p*. De la misma manera, el deseo de *p* responde a la cuestión de si *p* es deseable; la intención de hacer *p* responde a la cuestión de si *p* vale la pena hacerse; el miedo a *p* responde a la cuestión de si *p* representa una amenaza, y así sucesivamente. Otras actitudes, sin embargo, no responden a razones justificativas: en el mejor de los casos, podemos ofrecer razones explicativas sobre por qué tenemos hambre, nos duele la cabeza, o sentimos frío. En este sentido, recordar, imaginar o suponer no requieren, habitualmente, de justificación, y por eso tenemos mayor control sobre la ejecución de estas actividades mentales que sobre creencias o intenciones.

Por otra parte, no todas las actitudes que podrían ser susceptibles a razones lo son de hecho: podemos tener creencias delirantes, intenciones obsesivo-compulsivas, o miedos fóbicos. Estas actitudes se nos presentan alienadas, y como hemos señalado anteriormente, sólo las podemos conocer desde una posición teórica, de tercera persona. Las descubrimos, no las decidimos: no depende de

nosotros el llegar a tenerlas, mantenerlas o descartarlas, o al menos no en el sentido relevante para la atribución de responsabilidad.

¿En qué sentido dependen del agente las actitudes no alienadas, aquellas que sí responden a razones justificativas? Moran argumenta que podemos adoptar dos tipos de responsabilidad: *interna* y *externa* (Moran, 2001, pp. 117–118). Mientras que la responsabilidad interna requiere autoría en el sentido antes mencionado, la responsabilidad externa es la que adquirimos cuando manipulamos nuestras actitudes de manera no deliberativa. En lugar de preguntarnos si p es el caso, actuamos sobre el mundo, o sobre nosotros mismos, para forzar la aparición de esa creencia. Puedo manipular mi creencia de que esta mesa es azul pintándola de azul, tomar una bebida alcohólica para comportarme de manera más extrovertida, o asistir a terapia para crear un hábito positivo:

So, there is a sense in which a person can be active with respect to his sensations when he manipulates himself in one way or another to produce some feeling. And a similar stance is possible with respect to one's attitudes. In various cases a person may work to produce in himself various desires, beliefs or emotional responses, either by training, mental discipline, the cooperation of friends, or by hurling himself into a situation that will force a certain response on him. (Moran, 2002, p. 199)

Por el contrario, la responsabilidad interna es exclusiva de primera persona; es la que el agente tiene sobre sus actitudes en tanto que éstas expresan su criterio de evaluación. La responsabilidad interna del agente sobre estas actitudes es la de hacerlas responder a lo que para él son las razones que las justifican. Así, debo ajustar mis creencias a la evidencia que tengo disponible, mis deseos a lo que considero deseable, mis miedos a lo que considero temible, y así con el resto de actitudes sensibles al juicio.

Hieronymi (2008) desarrolla la distinción de Moran entre

responsabilidad interna y externa, y la reformula en términos del tipo de control que el agente tiene sobre sus actitudes. Por un lado, el control manipulativo, que correspondería a la responsabilidad externa, consiste en realizar alguna acción (física o mental) que cause un cambio de actitud. Por otro lado, en control evaluativo (por el que asumimos responsabilidad interna) consiste en asumir un compromiso hacia el contenido de las propias actitudes mediante la evaluación de las razones a favor de éstas:

Because these attitudes embody our take on the world, on what is or is not true or important or worthwhile in it, we control them by thinking about the world, about what is or is not true or important or worthwhile in it. Because our minds change as our take on the world changes [...] we can be said to be “in control” of our commitment-constituted attitudes. (Hieronymi, 2008, pp. 370–371)

Por lo tanto, la autoría, en tanto que expresión de los compromisos evaluativos del agente, depende directamente de su capacidad para evaluar los hechos del mundo, y convertirlos en sus razones, las que aduce para justificar sus actitudes. La responsabilidad es por tanto la obligación racional de relacionarse con los estados mentales propios de manera que éstos respondan a nuestras propias razones. Dicho de otra manera: nuestra obligación racional, en tanto que agentes, es formar y mantener actitudes de las que seamos internamente responsables. Hacer que las actitudes respondan a las propias razones es un proceso activo, agencial. Nuestra obligación racional es ser *autores* de nuestras actitudes, y no meros espectadores.

La posición de Moran presenta, a mi juicio, tres problemas. En primer lugar, la concepción externalista de Moran con respecto a las razones justificativas plantea problemas para la tesis de que el agente tiene un papel activo en la formación de sus estados mentales. En segundo lugar, la relación constitutiva entre razonamiento y responsabilidad que postula Moran no es ni necesaria ni suficiente. Por

último, la teoría de Moran no parece poder acomodar una intuición extendida acerca tanto de la responsabilidad como de la autoría: que éstas se dan en grados.

2. Concepción de las razones

Moran defiende una postura externalista con respecto a las razones, según la cual éstas son hechos que tienen un poder justificativo que es independiente de cuáles sean las creencias del sujeto:

What provides a reason for betting on a certain horse is the fact that it will win, or evidence for that fact, not one's belief, however strong, that it will win. And that fact would constitute a reason for the bet, even in the absence of the corresponding belief [...] A person's own attitude toward how his beliefs and desires relate to his action is not "subjectivist", or simply a matter of good fit among the states themselves. His belief is not for him a psychological datum that could even in principle justify his behavior purely in its role as a psychological state. (Moran, 2001, pp. 128–9)

Nuestra responsabilidad en tanto que agentes racionales consiste, según Moran, en hacer que nuestros estados mentales y acciones respondan apropiadamente a las razones que justifican estas actitudes y acciones. Sin embargo, la expresión "responder a razones" es susceptible de una doble interpretación, que se corresponde con la distinción clásica entre razones externas e internas. Por un lado, un agente que no responde a las razones que él cree que hay es considerado irracional. Si yo creo que p es el caso, y creo que de p se sigue necesariamente q , y al mismo tiempo creo que q no es el caso, estoy violando una norma de coherencia entre mis creencias, independientemente de que p o q sean o no el caso. Supongamos, siguiendo el famoso ejemplo de Williams (1979), que tengo sed, y que creo que lo que tengo delante de mí es un vaso de gin-tonic. En realidad, yo estoy equivocada: el vaso contiene gasolina en lugar de

ginebra. Tengo entonces razones internas para beber de ese vaso (ya que creo que el vaso contiene gin-tonic, y deseo beber gin-tonic), pero no tengo razones externas, ya que no hay ningún hecho en el mundo que justifique que beba de ese vaso. Puede decirse que, si no bebo, no estoy respondiendo adecuadamente a mis razones, y por lo tanto, no estoy siendo racional, en un sentido subjetivo.

Por otro lado, “responder adecuadamente a razones” implica, desde una posición externalista, responder adecuadamente a hechos del mundo cuya capacidad justificativa es independiente al agente. Volviendo al ejemplo de la gasolina, la postura externalista afirma que yo no tengo ninguna razón para beber lo que contenga ese vaso; en todo caso, tendría una razón “aparente” (Parfit, 2011, pp. 34–35). Puede ser racional para mí beber de ese vaso, pero eso no hace que yo tenga una razón justificativa para hacerlo. Mientras que la racionalidad es subjetiva, argumenta Parfit, y depende de la relación que nuestras actitudes guardan entre sí, las razones normativas son objetivas, ya que dependen de los hechos del mundo, no de nuestras actitudes mentales hacia estos hechos.

¿Cómo interpretar la postura de Moran? Si bien los hechos del mundo son razones independientemente de las creencias de los agentes, lo relevante es la relación del agente con el contenido de su creencia, es decir, con el hecho que es supuesto al tener una creencia acerca de éste. Si se me preguntara por qué bebo del vaso del ejemplo anterior, no respondería aludiendo a mi creencia (bebo porque creo que el vaso contiene gin-tonic), sino a su contenido: bebo porque el vaso contiene gin-tonic.

Así, a pesar de que las razones sean hechos cuya fuerza justificativa es independiente al agente, algo sólo es una razón en tanto que contenido de una creencia; la razón no es la creencia en sí, sino el hecho al que refiere:

Answering a deliberative question about one's belief regarding P is a matter of making up one's mind with respect to P, and here one's attention will be focused on the reasons for or against P, and not on facts about oneself. If it is possible for a person to answer a deliberative question about his belief at all, this involves assuming an authority over, and a responsibility for, what his belief actually is. Thus a person able to exercise this capacity is in a position to declare what his belief is by reflection on the reasons in favor of that belief, rather than by examination of the psychological evidence. (Moran, 2004, p. 425).

Podemos entonces entender la responsabilidad sobre las actitudes mentales como el resultado de un acto que, en realidad, es doble. Primero, el agente identifica el elenco de razones; luego, ajusta sus actitudes a estas razones. Siguiendo a Moran, el segundo acto es relativamente automático, ya que para zanjar la cuestión acerca de si creer que p, el agente atiende a sus razones, a los hechos que justifican la verdad de p (o a las razones para hacer F si trata de decidir si hacer F).

El problema, entonces, concierne a la responsabilidad que se le atribuye al agente sobre sus propias razones, ya que para explicar en qué sentido sus actitudes dependen de él habría que explicar primero en qué sentido sus razones dependen de él. Siguiendo la terminología de Hieronymi (2006), ¿tiene el agente control evaluativo sobre sus razones? Si lo tuviera, el argumento de Moran parecería ser circular, como señala Moya (2006): no explica el conocimiento que tenemos sobre los propios estados mentales, sino que lo presupone. Conocer las propias actitudes depende de conocer las propias razones, que a su vez se manifiestan en creencias. Si tenemos control evaluativo sobre estas creencias, entonces debemos también conocer las razones para tener estas creencias. Es necesario un mecanismo conceptual para detener

este regreso al infinito⁴.

De no tener control evaluativo, por el contrario, las creencias en cuyo objeto reconoce el agente una razón serían creencias alienadas, y las actitudes derivadas de estas razones estarían asimismo alienadas. La responsabilidad del agente sería solamente externa. Aquí se presentan dos problemas. Primero, las creencias alienadas pueden formar parte del proceso deliberativo. Segundo, las razones que no atañen al contenido de la actitud son ignoradas.

En primer lugar, una creencia alienada puede formar parte del proceso de razonamiento, desempeñando el papel de razón justificativa. Por ejemplo, supongamos que tengo fobia a los perros. Aunque sé que no tengo razones para temerlos, me aterroriza la posibilidad de encontrarme con un perro suelto. Y por ello, prefiero no pasear por los parques donde sé que las probabilidades de encuentro son altas. Así, mi creencia de que los perros son peligrosos no responde a mis razones (es decir, las razones que creo que hay), pero mi decisión de evitar ciertos parques de la ciudad, sí. A pesar de no ser ella misma una actitud alienada, mi decisión responde a razones que yo misma no reconozco como buenas razones.

En este punto, un atribucionista más afín a la propuesta de

⁴ La postura de Moran con respecto a la manera que tenemos de conocer nuestras razones admite varias interpretaciones. Una de ellas, que a mi juicio es la más plausible, es la apuntada por Moya: conocemos nuestras razones cuando tenemos algún tipo de creencia acerca de los hechos que consideramos razones. Sin embargo, como señala Fricke (2013), una interpretación alternativa podría salvaguardar el argumento de Moran de las acusaciones de circularidad y de regreso al infinito. Fricke sugiere que no es necesaria una creencia de segundo orden (“creo que p”) para tomar p como una razón. Discutir las dificultades de este enfoque, sin embargo, excede el ámbito de este trabajo.

Frankfurt (1971, 1987) podría argumentar que, aunque sean mis creencias alienadas las que me proveen de razones para formar otras creencias, puedo ser responsable de éstas si me identifico con ellas. Para Frankfurt, la identificación requiere de una actitud volitiva de segundo orden acerca de cómo resolver los conflictos entre nuestros deseos (o nuestras creencias) de primer orden⁵. Sin embargo, Moran toma la ruta contraria: “[W]hen a person does reach a (thought-dependent) desire through a process of practical reasoning he assumes a kind of responsibility for it, and he identifies with it” (Moran, 2002, p. 202); a lo que Frankfurt le responde: “The relationship dependence goes the other way around. It is identification that indispensability constitutes the source and the ground of reasons” (Frankfurt, 2002, p. 219). Así, Moran se distancia de la posición frankfurtiana, situando el control que el agente ejerce sobre sus actitudes mediante la deliberación, entendida como el proceso de responder a la pregunta sobre si algo es el caso, o sobre qué hacer.

En segundo lugar, al privilegiar el Principio de Transparencia para explicar la autoría, Moran sólo atiende a las razones a favor y en contra del contenido de las actitudes. La distinción entre *razones de estado* y *razones de contenido* (Parfit, 2002) suele aducirse para explicar por qué la creencia sólo responde a la evidencia, y no a los beneficios de poseer dicha actitud: el hecho de que nos ofrezcan una gran cantidad de dinero si creemos que p no es una razón para creer que p, o al menos no una a

⁵ El modelo jerárquico frankfurtiano, por otra parte, está sujeto a otros problemas (que Frankfurt reconoce ya en su artículo de 1971), tales como la posibilidad de un regreso al infinito en los niveles de actitudes (si puede haber de segundo orden, puede haber de tercer orden, y así sucesivamente), y la arbitrariedad del orden de la volición desde el cual el agente evalúa sus deseos de orden inferior (para un análisis sistemático de estas dos críticas, véase Ekstrom, 2005).

la que podamos responder simplemente formando esa creencia⁶. Sin embargo, tanto otras actitudes doxásticas, como la aceptación, como las intenciones suelen considerarse más permeables a este tipo de razones.

Así, puedo tener razones para aceptar una proposición, o para tomar una decisión práctica, que poco tienen que ver con lo que es aceptado o decidido. Por ejemplo, por razones de simplicidad o precisión, prefiero una explicación a otra rival (Harman, 2004). El hecho de que A sea más simple que B no es evidencia a favor de A, ya que no cuenta a favor de su verdad. De la misma manera, si decido pedir pasta en lugar de arroz en un restaurante lanzando una moneda al aire (ya que las alternativas me son absolutamente indiferentes), respondo a razones prácticas que no tienen que ver con la deseabilidad de los platos. En estos dos casos, mis actitudes responden a mis razones, pero no a las razones que atañen al contenido de mis actitudes. La formación de estas actitudes no respeta el Principio de Transparencia; ahora bien, ¿soy por ello menos responsable de estas actitudes?

En conclusión, para mantener una teoría de la responsabilidad basada en las razones, sería necesario elucidar con más detalle la transmisión de la autoría, y por tanto de la responsabilidad, a través de los mecanismos evaluativos, desde las razones hasta las actitudes mentales y las acciones, así como aclarar cuál es la relación racional entre hechos del mundo, justificación y creencias que fundamenta un concepto de razón normativa adecuado para el concepto de autoría.

3. Relación actitudes – razonamiento

⁶ Algunos autores como Reisner (2009) o Schroeder (2012) han argumentado en contra de esta distinción. Aunque simpatizo con sus críticas, me centraré en el caso de la actitud de *aceptación*, por razones de simplicidad.

Según Moran, es imprescindible para la autoría que el agente adopte una actitud deliberativa. Sin embargo, las creencias, intenciones y deseos no se forman, habitualmente, mediante el razonamiento explícito. Moran admite que, de hecho, nuestras actitudes mentales no suelen ser la conclusión de un proceso deliberativo; relajando el requisito deliberativo, Moran establece que, para la relación de autoría, basta con que estas actitudes *puedan ser* la conclusión de tal proceso. Es decir, mientras que nuestras actitudes no tienen por qué estar necesariamente formadas mediante un proceso de deliberación, sí han de ser la conclusión posible del razonamiento teórico (en el caso de las creencias y otras actitudes doxásticas) o práctico (en el caso de los deseos y otras actitudes prácticas) (Moran, 2001, p. 116).

Ahora bien, es difícil entender el sentido de “posibilidad” con respecto a la justificación de una actitud. Podría entenderse que aunque la actitud no haya sido formada en un proceso de razonamiento cuya conclusión fuera la actitud en cuestión, el agente ha de poder justificar la actitud. Por ejemplo, al ver un perro que muestra sus dientes, puedo sentir miedo, pero no como resultado de un proceso de deliberación, sino de manera espontánea, al responder al hecho de que un perro esté en posición amenazante. Puedo, además, justificar mi miedo: hay un perro en posición amenazante. Bajo esta interpretación, la relación de autoría requiere que el agente sea capaz de dar razones que justifiquen la actitud o la acción.

Sin embargo, ser capaz de justificar una actitud no implica que la razón por la que mantenemos esta actitud sea la que la justifica. El hecho de que nuestros deseos, creencias y demás actitudes sean susceptibles de ser justificados racionalmente no implica que tengamos esas actitudes en base a consideraciones racionales. Podemos considerar que hay muy buenas razones para hacer algo, y sin embargo

hacerlo por algún motivo banal, o por un motivo con el que no nos identifiquemos. Una persona que sufre de un trastorno obsesivo compulsivo y siente la necesidad de lavarse las manos puede encontrar buenas razones para lavárselas (por ejemplo, que va a comer algo con las manos), pero no quiere decir que sean éstas razones las que motiven la acción; de hecho, es su pánico a la suciedad lo que motiva su intención y su acción. Dicho de otro modo, podemos tener razones motivacionales que no se corresponden con nuestras razones justificativas.

Hay una segunda manera, más restrictiva, de interpretar en qué consiste ser la conclusión posible del razonamiento, según la cual no sólo el agente ha de poder ofrecer razones que justifiquen su actitud, sino que mantiene la actitud por esas mismas razones. Sin embargo, como argumentaré a continuación, esta condición no es ni necesaria ni suficiente para la autoría: podemos ser responsables de estados que no son la conclusión posible de nuestro razonamiento, así como no serlo de estados que son, de hecho, la conclusión de nuestro razonamiento.

En primer lugar, podemos tener actitudes en contra de nuestras razones, y ser sin embargo autores de estas razones, y por tanto responsables de ellas. La akrasía consiste, precisamente, en tener actitudes intencionales o realizar acciones en contra de las mejores razones⁷. No obstante, no todas las actitudes akráticas están alienadas, ya que no siempre el agente pierde el control, como sí sucede con el ejemplo del ludópata akrático que Moran toma de Sartre. Este ludópata decide dejar el juego, pero sabe que es muy probable que actúe en contra de esta decisión. Moran señala que el ludópata sólo puede averiguar cuál es su grado de determinación reflexionando con actitud

⁷ Sigo aquí el concepto de akrasía desarrollado por Holton (2009); la debilidad de la voluntad, por el contrario, consistiría en revisar precipitadamente y sin justificación intenciones previas.

teórica, de tercera persona, sobre cómo se ha comportado en el pasado, y no adoptando una actitud práctica hacia su decisión, reflexionando sobre las razones que tiene para abandonar el juego (Moran, 2001, p. 94). El ludópata sabe que carece de la voluntad necesaria para hacer que sus decisiones controlen su comportamiento. La adicción y la compulsión generan actitudes y comportamientos akráticos en tanto que socavan las capacidades volitivas del agente, ya sea total o parcialmente⁸; pero un agente puede tener actitudes akráticas cuando sus razones normativas no acompañan a sus motivaciones, sin que sus capacidades agenciales se vean afectadas, y sin eludir la responsabilidad sobre sus actitudes y acciones. Por ejemplo, tras deliberar acerca de si debería irme de vacaciones a París, concluyo que no debería hacerlo, puesto que para ello necesitaría sacrificar unos ahorros que tenía reservados para emergencias. Sin embargo, decido ir a París siguiendo otras razones cuyo poder justificativo es menor, pero que me motivan en mayor grado. Si finalmente necesitara en el futuro de los ahorros que gasté, pensaría igualmente que no debería haberme ido a París de vacaciones (aunque si la experiencia fue muy positiva, podría racionalizar la acción, pensando que mereció la pena), pero no pensaría que mis capacidades volitivas estuviesen socavadas. Decidí actuar en contra de mis razones, y soy responsable y autora de mi decisión y mis acciones.

En segundo lugar, el que un estado mental sea el resultado de un proceso de deliberación no basta para que el agente se considere autor de este estado. Como señalábamos en el apartado anterior, las razones del agente pueden serle ajenas; por ejemplo, tener miedo fóbico a los perros hace que considere que el hecho de que hay perros

⁸ Véase, para un análisis de cómo la adicción y la compulsión afectan a la responsabilidad y a las capacidades volitivas, Gideon Yaffe (2001) y Watson (2004).

en el parque es una razón para evitar ir al parque. Aquí, el agente es consciente de que su miedo es irracional, y por lo tanto sabe que, de no padecer cinofobia, no tendría ninguna razón para no ir al parque. La decisión de no ir al parque, no obstante, es la conclusión del razonamiento del agente, pero al estar sus razones alienadas, también lo está la conclusión.

Por lo tanto, ni la interpretación más restrictiva de la tesis de la relación necesaria entre autoría y razonamiento, ni la menos restrictiva, reflejan adecuadamente la interacción entre razones justificativas y motivaciones que da lugar a las actitudes mentales.

4. Grados de responsabilidad

La última observación a la teoría de Moran (y que también puede aplicarse al desarrollo de ésta por parte de Hieronymi en su artículo de 2008) está dirigida a señalar una carencia, relativa al problema de los grados de agencia y de responsabilidad interna. Intuitivamente, la responsabilidad es un concepto que admite grados. De la misma manera, no todas nuestras acciones muestran el mismo grado de autoría. Sin embargo, esta intuición no es fácilmente acomodable en una teoría que hace depender la responsabilidad del control evaluativo del agente. El tipo de gradación que sería relevante incorporar está relacionado con la supuesta diferenciación estricta entre actitudes alienadas y no alienadas. Podemos imaginar ejemplos (como el expuesto sobre la fobia a los perros) en los que, intuitivamente, existe algún grado de alienación, pero no una exención plena de responsabilidad. El caso del adicto akrático, cuyo deseo de jugar le está claramente alienado, ilustra eficientemente el significado de poseer una actitud sobre la que desearíamos poder ejercer control, pero no podemos y por ello sólo podemos aspirar a ser meros espectadores, o a tratar de controlarla de manera externa (yendo a terapia, por ejemplo).

Sin embargo, es más frecuente (a mi juicio) que los agentes tengan actitudes sobre las que no disponen de un control pleno, pero de las que son responsables internamente, no sólo de manera externa.

Sería posible abordar la cuestión de los grados de responsabilidad sin abandonar una teoría basada en las razones si se adoptara un marco diferente para explicar cómo el agente puede lograr que sus actitudes respondan a las razones que considera mejores. La propuesta de Fischer y Ravizza (2000)⁹ consiste en explicar la competencia normativa, es decir, la capacidad de responder a razones, mediante la introducción de mecanismos de responsividad a estas razones, que involucran capacidades cognitivas y volitivas. Así, el grado de responsabilidad de un agente por una actitud dependerá de hasta qué punto ha sido capaz de reconocer y reaccionar frente a sus razones para mantener esta actitud. Por ejemplo, Goldie (2004) señala cuatro tipos de factores que alteran el comportamiento influyendo causalmente en la conducta del agente, desviándolo de lo que sería su respuesta habitual: estados alterados (como la influencia de drogas, la falta continuada de sueño o la fiebre), ciertas emociones (como los celos o la ira), algunos estados más o menos duraderos (como la depresión, la irritabilidad o el estrés), y por último, factores situacionales¹⁰. Podría interpretarse que estos cuatro factores afectan al comportamiento del agente haciéndolo menos responsivo a razones, de manera que los mecanismos evaluativos que suele emplear para reconocer las razones que justifican sus acciones y sus actitudes se ven alterados. Una persona que padece

⁹ Aunque aquí sólo señalo la propuesta de Fischer y Ravizza, existen otras propuestas afines; véase Nelkin y Brink (2013) para un análisis reciente de las diversas teorías de la responsabilidad basadas en la responsividad a razones.

¹⁰ El trabajo de Goldie se centra en el experimento de la cárcel de Stanford; Nelkin (2005) examina diversos experimentos situacionistas, y analiza su relación con las atribuciones de responsabilidad.

un ataque de celos forma creencias que no responden a las razones que consideraría si no se sintiese celosa, y por lo tanto su autoría y responsabilidad sobre estas creencias es menor que si las formase en un estado emocional no alterado.

Mi intención no es entrar a valorar las propuestas de Fischer y Ravizza y de Goldie, sino señalar que es posible desarrollar la idea de que la responsabilidad requiere de la capacidad de evaluar a razones y de responder a ellas, sin comprometerse con la dicotomía entre estados alienados y no alienados. El concepto de responsabilidad es una pieza clave en la manera que tenemos de pensar la normatividad. Elucidar este concepto no es sólo relevante para las teorías éticas, sino también para las teorías de la agencia.

Conclusión

Las teorías atribucionistas de la responsabilidad capturan una intuición filosófica muy expandida: que debe haber algo del autor en la obra para que podamos atribuirla, y que en virtud de ese algo, la obra “le pertenece”. Pero también se enfrentan al problema de la esencialidad del “verdadero yo”: cuando las acciones o actitudes expresan el verdadero ser, ¿qué es lo que expresan? ¿En qué sentido las razones del agente definen o retratan su carácter moral?

La teoría constitutivista acerca del autoconocimiento presentada por Moran resulta atractiva al evitar comprometerse con la imagen cartesiana del agente como espectador del teatro que sucede en su mente¹¹. Conocemos nuestros estados mentales porque los decidimos; y al decidirlos, nos responsabilizamos de ellos. Esta relación

¹¹ Véase Lazos (2008).

entre responsabilidad, razones y autoría necesita, como he tratado de mostrar, desarrollar más la cuestión de qué sean las razones, y de cómo el agente las reconoce y se las apropia, si queremos fundamentar en ellas una teoría de la responsabilidad.

REFERENCIAS BIBLIOGRÁFICAS

- BUSS, S. & OVERTON, L. (eds.). *Contours of agency: Essays on themes from Harry Frankfurt*. Cambridge, Mass: MIT Press. 2002.
- DEL CORRAL, M. “el problema de las razones inadecuadas”. *Factótum: Revista de Filosofía*, 11, pp. 103–11, 2014.
- EGONSSON, D.; JOSEFSSON, J.; PETERSSON; B. RONNOW-RASMUSSEN, T. & PERSSON, I. (eds.). *Exploring practical philosophy: From action to values*. Burlington: Ashgate Publishing, 2002.
- EKSTROM, L. W. “Alienation, autonomy, and the self”. *Midwest Studies in Philosophy*, 29(1), pp. 45–67.
- FISCHER, J. M., RAVIZZA, M. *Responsibility and control: A theory of moral responsibility*. Cambridge, Mass: Cambridge University Press, 2002.
- FRANKFURT, H. “Freedom of the Will and the Concept of a Person”. *The Journal of Philosophy*, 68(1), pp. 5–20, 1971.
- FRANKFURT, H. “Identification and wholeheartedness”. In F. Shoeman (ed.) (1987), pp. 27–45.
- FRANKFURT, H. “Reply to Richard Moran”. In S. Buss and L.
- Manuscrito – Rev. Int. Fil., Campinas, v. 38, n.1, pp. 207-230, jan.-jun. 2015.

- Overton (eds.) (2002), pp. 218–225.
- FRICKE, M.F. “First Person Authority and Knowledge of One’s Own Actions”. *Crítica: Revista Hispanoamericana de Filosofía* 45 (134), pp. 3–16, 2013.
- GIDEON YAFFE. “Recent Work on Addiction and Responsible Agency”. *Philosophy & Public Affairs*, 30(2), pp. 178–221, 2001.
- GOLDIE, P. “What people will do: personality and prediction”. *Richmond Journal of Philosophy*, 7, pp. 11–18, 2004.
- HARMAN, G. “Practical Aspects of Theoretical Reasoning”. In A. R. Mele & P. Rawling (eds.) (2004), pp. 45–56.
- HARRISON, R. (ed.). *Rational Action: Studies in Philosophy and Social Sciences*. Cambridge: Cambridge University Press, 1979.
- HIERONYMI, P. “Controlling attitudes”. *Pacific Philosophical Quarterly*, 87(1), pp. 45–74, 2006.
- HIERONYMI, P. “Responsibility for believing”. *Synthese*, 161(3), pp. 357–373, 2008.
- HOLTON, R. *Willing, wanting, waiting*. Oxford: Oxford University Press, 2009.
- LAZOS, E. “Autoconocimiento: Una Idea Tensa.” *Diánoia* 53 (61): pp. 169–88, 2008.
- MELE, A.R., RAWLING, P. (eds.). *The Oxford Handbook of Rationality*. Oxford, New York: Oxford University Press, 2004.
- MORAN, R. “Self-Knowledge: Discovery, Resolution, and Undoing”. *European Journal of Philosophy*, 5(2), pp. 141–161, 1997.
- MORAN, R. *Authority and estrangement: An essay on self-knowledge*. Princeton, New Jersey: Princeton University Press, 2001.

- MORAN, R. "Frankfurt on identification: Ambiguities of activity in mental life". In S. Buss & L. Overton (eds.) (2002), pp. 189–217.
- MORAN, R. "Responses to O'Brien and Shoemaker". *European Journal of Philosophy*, 11(3), pp. 402–419, 2003.
- MORAN, R. "Précis of Authority and Estrangement: An Essay on Self-Knowledge". *Philosophy and Phenomenological Research*, 69(2), pp. 423–426, 2004.
- MOYA, C. J. 2006. "Moran on Self-Knowledge, Agency and Responsibility." *Crítica: Revista Hispanoamericana de Filosofía* 38 (114): 3–20.
- NELKIN, D. K. "Freedom, responsibility and the challenge of situationism". *Midwest Studies in Philosophy*, 29(1), 1pp. 81–206, 2005.
- NELKIN, D. K., BRINK, D. O. (2013). "Fairness and the Architecture of Responsibility". In D. Shoemaker (ed.) (2013), pp. 284–314.
- O'BRIEN, L. *Self-knowing agents*. Oxford; New York: Oxford University Press, 2007.
- PARFIT, D. "Rationality and reasons". In D. Egonsson, J. Josefsson, B. Petersson, T. Ronnow-Rasmussen, and I. Persson (eds.) (2002), pp. 17–39.
- PARFIT, D. *On what matters*. Oxford: Oxford University Press. 2011.
- PETTIR, P., SMITH, M. "Freedom in belief and desire". *The Journal of Philosophy*, 93(9), pp. 429–449, 1996.
- REISNER, A. The possibility of pragmatic reasons for belief and the wrong kind of reasons problem. *Philosophical Studies*, 145(2),

pp. 257–272, 2009.

SCANLON, T. *What we owe to each other*. Cambridge, Mass: Belknap Press, 1998.

SHOEMAN, F. (ed.). *Responsibility, Character and the Emotions*. New York: Cambridge University Press, 1987.

SHOEMAKER, D. (ed.). *Oxford Studies in Agency and Responsibility*. Oxford: Oxford University Press, 2013.

SCHROEDER, M. “The Ubiquity of State-Given Reasons”. *Ethics*, 122(3), pp. 457–488, 2012.

VARGAS, M. “Reasons and real selves”. *Ideas y Valores*, 58(141), pp. 67–84, 2009.

WATSON, G. “Disordered appetites: Addiction, compulsion, and dependence”. In *Agency and Answerability* (pp. 59–88). Oxford, New York: Oxford University Press, 2004.

WILLIAMS, B. A. O. “Internal and External Reasons”. In R. Harrison (ed.) (1979) pp. 17–28.