

¿Pueden los animales ser morales?

Mark Rowlands

University of Miami

mrowlands@mail.as.miami.edu

Can Animals be Moral?

RESUMEN: La distinción dicotómica entre agentes y pacientes morales está ampliamente reconocida y bien asentada en filosofía moral. Los animales, en el caso de ser considerados como poseedores de estatuto moral alguno, son, casi invariablemente, reconocidos como pacientes morales más que como agentes. La tesis principal de este artículo sostiene que hay una tercera opción: mientras los animales son pacientes morales, en lugar de agentes morales, pueden ser también sujetos morales, donde: X es un sujeto moral si y solamente si X está, por lo menos algunas veces, motivado a actuar por consideraciones morales. Mediante una serie de argumentos, el artículo trata de mostrar que no hay obstáculo lógico o conceptual alguno que nos impida pensar en los animales como motivados por consideraciones morales, específicamente por emociones con un contenido moral identificable.

ABSTRACT: The dichotomy between moral agents and moral patients is a recognized and well established one in moral philosophy. Animals, if they are thought of as having moral standing at all, are, almost invariably, regarded as moral patients rather than moral agents. The main thesis of this paper is that there is a third option: while animals are indeed moral patients, and not moral agents, they can also be moral subjects, where: X is a moral subject if and only if X is, at least sometimes, motivated to act by moral considerations. By means of a series of arguments, the paper attempts to show that there are no logical or conceptual obstacles to thinking of animals as motivated by moral considerations, specifically by emotions that have identifiable moral content.

PALABRAS-CLAVE: animales no humanos, teoría de la acción, agencia moral, subjetividad moral, emociones, Aristóteles, Wittgenstein

KEYWORDS: non human animals, action theory, moral agency, moral subjecthood, emotions, Aristotle, Wittgenstein

ISSN 1989-7022

FILEMATA año 4 (2012), n° 9, 1-32

1. Gracia, de las Virtudes

Eleonora, la matriarca de su familia, se muere y es incapaz de tenerse en pie. Gracia le ayuda cuidadosamente a ponerse de nuevo en pie. Intenta que Eleonora camine, empujándola con sumo cuidado. Pero Eleonora vuelve a caer. Gracia parece profundamente afligida y chilla con fuerza. Trata de levantar de nuevo a Eleonora, pero sus esfuerzos son en vano. Gracia permanece junto a Eleonora durante una hora;

Reconocimientos:

Traducción de Mikel Torres Aldave (UPV/EHU).

Texto perteneciente a la conferencia plenaria impartida el 2 de Junio de 2011 en la 17ª SEMANA DE ÉTICA Y FILOSOFÍA POLÍTICA "Nosotros y los otros", celebrada en la Facultad de Filosofía y Ciencias de la Educación de la Universidad del País Vasco/Euskal Herriko Unibertsitatea (UPV/EHU) en Donostia – San Sebastián, entre el 1 y el 3 de junio de 2011. El texto es un resumen de las principales ideas y argumentos que el autor presentará en su libro *Can Animals be Moral?*, a publicar en 2012 por Oxford University Press. Las referencias y notas al pie del original se han omitido para simplificar el texto e incentivar la lectura del libro.



Received: 06-03-2012

Accepted: 13-03-2012

mientras, va anocheciendo. Si Gracia fuera humana, no tendríamos ningún reparo en atribuirle cierto tipo de emociones como compasión o simpatía. Pero Gracia no es humana, tampoco Eleonora lo es. Eleonora es la matriarca de la familia de elefantes Primeras Damas. Gracia es un miembro más joven de otra familia de elefantes, la familia Virtudes. Este es el tipo de caso que se suele citar como evidencia de que el comportamiento de los animales no humanos (en lo sucesivo “animales”) puede estar motivado por consideraciones morales. En este artículo no podré estudiar el amplio y creciente conjunto de investigaciones empíricas sobre el que se sustenta una afirmación de este tipo. Este trabajo se sitúa en el fondo de la cuestión – quiere ser el trampolín para una discusión más abstracta y conceptual. Utilizo este caso como ejemplo representativo de esta investigación, así como un camino mediante el cual organizar la discusión.

En este artículo, defenderé la tesis de que Gracia *podría* ser un sujeto moral. Argumentaré que no hay obstáculo lógico o conceptual alguno que nos impida pensar en Gracia como motivada por consideraciones morales – específicamente, por emociones con un contenido moral identificable. La motivación real de Gracia es, por supuesto, una cuestión empírica, y no pretendo adoptar al respecto postura alguna. Simplemente defiendiendo la posibilidad de un juicio de este tipo: es posible que los animales estén motivados por estados emocionales como la compasión, la simpatía, y la tolerancia, y también por sus contrapuestos negativos como los celos, la malicia, y el resentimiento. Todas estas son emociones con contenido moral – emociones *con carga moral*, como las denominaré.

Alguien puede pensar que por el simple hecho de defender la posibilidad de un juicio de este tipo, estoy marcando las cartas a mi favor de manera inaceptable. En realidad, queda mucho trabajo por hacer. La atribución de motivación moral a los animales se considera de manera casi general – por los filósofos, al menos – imposible. En parte, los filósofos tienen buenas razones para ello. Pero, también en parte, están confundidos. Y, en parte, han caído bajo el hechizo de cierto tipo de pensamiento mágico. Pero antes de que pueda siquiera llegar a este punto, necesito presentar de forma algo más precisa la tesis principal de este artículo.

2. Emociones con carga moral

La tesis que defenderé a continuación apela al concepto de emoción con carga moral. En esta sección, explicaré lo que entiendo por esto. Desafortunadamente, definiendo una más que compleja teoría sobre la atribución de emociones que aquí solamente podré esbozar, no defender.

Las emociones, en el sentido en que las entiendo en este artículo, son estados intencionales. Tienen contenido en la forma en que los pensamientos, las creencias y los deseos lo tienen. Podemos atribuírselo por medio de una oración subordinada¹. Por ejemplo: A Smith le indigna *que* Jones le desairara. Que Jones le desairara es el contenido de la indignación de Smith. La atribución de estados emocionales a un sujeto, S, por medio de una proposición, *p*, no implica que S contemple que *p*, ni tampoco que sea capaz de contemplarla. El uso de *p* está garantizado si *p* implica – en sentido amplio, se aproximaría a ‘es cierto en un contexto dado’ – otra proposición que S contempla. La idea de que un animal pueda contemplar una proposición podría parecer inverosímil. Pero una proposición es, por supuesto, no una oración sino el contenido expresado por la oración. Si los animales pueden contemplar contenidos – si pueden pensar algo – pueden contemplar proposiciones.

Establecido este marco, existen dos modos en que la posesión de una emoción podría, digámoslo así, *fallar* – lo cual entenderé como, aproximadamente, análogo al hecho de que una creencia sea falsa. La categoría de fallos es conjuntiva: una emoción falla cuando está *equivocada* o cuando está *extraviada*. Smith está indignado porque cree que Jones le ha desairado, pero él está, de hecho, en un error. Jones no le desairó en ningún momento. Digamos que la indignación de Smith está, en este caso, *equivocada*. Una emoción está equivocada cuando viene predicada de una aserción factual que resulta ser falsa. Para que su indignación no esté equivocada, la proposición factual ‘Jones me desairó’ tiene que ser verdadera. Si no está equivocada, entonces la indignación de Smith convierte en verdadera esta proposición factual. Supongamos, ahora, que Jones efectivamente desairó a Smith. Sin embargo, ella estaba en su derecho de hacerlo – Smith estuvo insoportable durante su

último encuentro. Smith, podríamos decir, no merecía nada mejor por parte de Jones en este caso. La indignación de Smith, en este caso, está *extraviada*. La emoción está *extraviada*, en este caso, porque se basa en una asunción de derecho que es, de hecho, errónea. Así pues, para que una emoción, *E*, no esté *extraviada*, debe haber una proposición evaluativa, *p*, que tiene que ser cierta. No es necesario que el sujeto de *E* sea capaz de contemplar *p*. Pero la verdad de *p* es necesaria para que *E* no esté *extraviada*: la verdad de *p*, podríamos decir, *da sentido* a *E*. La indignación de Smith, si no está *extraviada*, convierte en verdadera la proposición 'Jones hizo mal al desairarme'. Que una emoción no esté equivocada implica la veracidad de la proposición factual. Que una emoción no esté *extraviada* implica la veracidad de la proposición evaluativa.

Pertrechados con estas ideas, podemos definir el concepto de emoción moralmente cargada como sigue:

Una emoción, *E*, está *moralmente cargada* si: (1) existe una proposición, *p*, que expresa una reivindicación moral, y (2) si *E* no está *extraviada*, entonces *p* es verdadera.

En este artículo, defenderé que los animales pueden ser *sujetos* morales en el sentido de que son sujetos con emociones moralmente cargadas. La categoría de sujeto moral es una categoría nueva que no se reduce a las categorías más familiares de agente y paciente moral.

3. Agentes, pacientes y sujetos

Los animales, en el caso de ser considerados como poseedores de estatuto moral alguno, son, casi invariablemente, reconocidos como *pacientes* morales más que como *agentes*, donde:

- (1) *X* es un *paciente* moral si y solamente si *X* es un objeto legítimo de consideración moral: es decir, aproximadamente, *X* es una entidad que tiene intereses que deberían ser considerados a la hora de tomar decisiones que le afecten.

- (2) X es un *agente* moral si y solamente si X es (a) moralmente responsable de sus motivos y acciones, y por tanto puede ser (b) moralmente evaluado por ellos (alabado o reprendido, en sentido amplio).

Nada en (1) y (2), por supuesto, descarta que el mismo individuo sea al mismo tiempo un agente y un paciente moral: la mayoría de seres humanos son ambas cosas.

En lo que concierne a humanos y animales, habitualmente se asume que esto es todo lo que la geografía lógica se extiende. Los animales pueden ser tanto agentes como pacientes morales, o ambas cosas al mismo tiempo. No son, se presupone normalmente, agentes morales. Por lo tanto, no pueden ser más que pacientes morales. Esta es la historia ortodoxa. Existen, no obstante, voces discordantes. Por ejemplo, David DeGrazia escribe:

Estos ejemplos apoyan la atribución de agencia moral – específicamente, acciones que manifiestan virtudes – en casos en los que las acciones no pueden ser plausiblemente interpretadas como instintivas o condicionadas. En cualquier concepción razonable de la agencia moral, algunos animales son agentes morales.

Steven Sapontzis se hace eco de esta afirmación y también Evelyn Pluhar, la cual escribe:

¿Realmente está tan claro, sin embargo, que la capacidad para la agencia moral no tiene precedente alguno en ninguna otra especie? Ciertas capacidades adicionales son requeridas para la agencia moral, incluyendo capacidades para la emoción, memoria y el comportamiento orientado a un fin. Como hemos visto, existen amplias evidencias para afirmar la presencia de dichas capacidades, aunque sea en un grado limitado, en algunos no humanos. De manera no sorprendente, entonces, han sido reunidas evidencias que indican que algunos no humanos son capaces de lo que denominaríamos comportamiento "moral" o "virtuoso".

Afirmaciones similares, aunque en las formas más variadas, pueden encontrarse en los trabajos de Vicki Hearne, Jeffrey Moussiaeff Masson, Susan McCarthy, Stephen Wise, Frans de Waal y Marc Bekoff. Incluso

Darwin afirmó que los animales pueden estar motivados por “sentimientos morales”. Tanto DeGrazia como Pluhar expresan su afirmación en el lenguaje de la agencia. Esto es, pienso, desafortunado. El concepto de agencia es inseparable del concepto de responsabilidad y, por tanto, de los conceptos de alabanza y reprensión. Si los animales son agentes morales, se sigue que tienen que ser responsables de sus actos y, por tanto, pueden ser alabados o reprendidos por lo que hacen. Hubo un tiempo en el que los intentos de los tribunales – tanto seculares como religiosos – de procesar (y subsecuentemente ejecutar) animales por crímenes cometidos no eran infrecuentes. Asumo que muy pocos recomendarían que volviésemos a esas prácticas. En el centro de esta reticencia se encuentra el pensamiento de que los animales no son responsables y, por tanto, no pueden ser culpables de lo que hacen. Si esto es correcto, entonces deberíamos resistirnos a su caracterización en términos de agencia moral.

Argumentaré que hay otra opción: mientras los animales son pacientes morales, en lugar de agentes morales, pueden ser también *sujetos* morales, donde:

- (3) X es un *sujeto* moral si y solamente si X está, por lo menos algunas veces, motivado a actuar por consideraciones morales.

El concepto de sujeto moral ha sido casi invariablemente unido con el de agente moral: decir que Smith está motivado a actuar por consideraciones morales se considera invariablemente como equivalente a decir que Smith es responsable de, y por tanto puede ser alabado o reprendido por, sus actos. En cierto sentido esta unión es extraña: las definiciones (2) y (3) dejan claro que estas afirmaciones son lógicamente diferentes. (2) es una afirmación sobre una evaluación; (3) es una afirmación sobre una motivación. La agencia moral y la subjetividad moral² deberían ser tan conceptualmente distintas como lo son los conceptos de evaluación y motivación. Y está razonablemente claro que, *en general*, estas son cosas completamente diferentes: la motivación de una acción es una cosa, la evaluación de la acción o la motivación otra diferente. Aún más, una evaluación lo es muchas veces *de* una motivación.

Supongamos, por ejemplo, que mi esposa, agobiada por la suciedad doméstica que acompaña a dos chicos jóvenes, un perro grande y un marido desaliñado, me ha hipnotizado (sin que yo lo sepa). Ahora, siempre que pronuncia la palabra "Rosebud", experimento el deseo irrefrenable de fregar el suelo. Este deseo es, parece, un estado motivacional, uno que combinado con los estados cognitivos relevantes (la creencia de que esto es una fregona, la creencia de que esto es un suelo, y así sucesivamente) resultará, *ceteris paribus*, en un cierto tipo de comportamiento por mi parte. Este pasar la fregona, sin embargo, no es algo por lo que pueda ser alabado o reprendido. *Ex hypothesi*, es el resultado de un estado motivacional que está fuera de mi control. En otras palabras, mi comportamiento está motivado por cierto estado, pero ese estado no es el tipo de cosa por la cual yo pueda ser moralmente evaluado. Si el determinismo extremo – de hecho, la generalización de un caso de este tipo – fuera cierto, entonces nadie podría ser moralmente evaluado por sus actos, pero de ahí no se seguiría que no fueran sujetos de estados motivacionales.

No obstante, hay razones persuasivas – persuasivas en el sentido de que casi todo el mundo ha sido persuadido por ellas – para suponer que esta distinción general entre motivación y evaluación no es aplicable en el caso moral: es decir, en el caso *específico* de la motivación y evaluación moral. Así pues, el punto de vista estándar afirma que no hay distinción alguna entre un agente y un sujeto moral. Las razones de esto no son difíciles de entender...

4. Normatividad y agencia

Imaginemos a alguien – Smith – para quien sus motivaciones siempre permanezcan ocultas. Esto no significa que Smith carezca de todo autoconocimiento – él es consciente de sus estados no motivacionales: es la facultad motivacional de su mente la que se le oculta. Esta parte de su mente es algo así como una caja negra: repleta de estados que guían el comportamiento de Smith de forma exitosa, pero a los cuales no tiene

acceso directo. Es posible que Smith se muestre algo perplejo al ver que reacciona de forma adecuada ante las exigencias circunstanciales del entorno. Cuando los autobuses van a toda velocidad hacia él, cuidadosamente da un paso atrás. No tiene ni idea de qué le motiva a evitar el autobús; pero lo evita en cualquier caso. Smith encuentra difícil de explicar gran parte de su comportamiento. Hay un sentido obvio en el que Smith está, podríamos decir, "a merced" de sus motivaciones. No tiene ni idea de qué le motiva a actuar en la forma en que lo hace y, por tanto, no tiene control alguno sobre sus motivaciones. A causa de ello, Smith podría no ser un agente en el sentido en que lo es un adulto humano normal: él es arrastrado en una u otra dirección por motivaciones que le están vedadas. Y si Smith no es un agente, entonces, *a fortiori*, no es un agente moral. No obstante, es todavía cierto que él se encuentra motivado a actuar de ciertas maneras, a pesar de que permanezca con visión ciega ante dichas motivaciones. Incluso aunque no sea un agente, Smith es, sin embargo, el sujeto de una motivación. ¿Puede también ser el sujeto de una motivación específicamente *moral*? El control de sus motivaciones podría ser un requisito que Smith debería cumplir para ser un agente moral. Pero, ¿por qué debería serlo para que Smith fuera considerado un sujeto moral?

La respuesta puede encontrarse en el dictum Kantiano según el cual *deber implica poder*. Como Smith está con visión ciega ante sus motivaciones, no tiene control alguno sobre ellas. Ellas no son, por lo tanto, el tipo de cosas ante las que Smith puede rendirse o resistirse. Pero si él no puede rendirse o resistirse ante sus motivaciones, no tiene sentido decir que *debería* rendirse o resistirse ante ellas. Las motivaciones de Smith, en este sentido, no tienen dimensión normativa alguna. No son el tipo de cosas que él *debería* aprobar o rechazar. Sus motivaciones no realizan *reivindicación normativa* alguna sobre Smith. Sin embargo, las motivaciones morales son precisamente el tipo de cosas que realizan reivindicaciones normativas sobre los sujetos. Las buenas motivaciones morales son aquellas que deberían ser aprobadas por los sujetos de las mismas; las malas motivaciones morales son aquellas que deberían ser rechazadas. Así pues, parece que estamos obligados a rechazar que las motivaciones de Smith sean motivaciones morales.

La idea de que las motivaciones no pueden clasificarse como morales a menos que el sujeto tenga control sobre ellas ha influido decisivamente en la mayoría de discusiones históricas influyentes en torno a la motivación moral. Ahora es evidente por qué esto debería ser así: el control sobre las propias motivaciones es una condición necesaria para que éstas tengan estatus normativo – es decir, para que realicen una reivindicación normativa sobre aquellos sujetos que las poseen.

5. Kant y Aristóteles: La condición de reflexión

Christine Korsgaard, una de las principales representantes contemporáneas del enfoque Kantiano en filosofía moral, escribe:

Kant creía que los seres humanos han desarrollado una forma específica de auto-conciencia, a saber, la habilidad de percibir, y por tanto pensar sobre, los fundamentos de nuestras creencias y acciones en tanto que fundamentos. Lo que quiero decir es: un animal que actúa por instinto es consciente de sus objetos de temor o deseo, y consciente de ellos como aterradores o deseables, y como objetos que deben ser evitados o que deben ser perseguidos. Ése es el fundamento de sus acciones. Pero un animal racional es, adicionalmente, consciente de que teme o desea esos objetos, y de que se muestra inclinado a actuar de cierta manera como resultado de ello. Esto es lo que quiero decir cuando hablo de ser consciente del fundamento como fundamento. Así que como seres racionales somos conscientes de los principios sobre los que nos basamos a la hora de actuar. Por eso mismo, tenemos la habilidad de preguntarnos a nosotros mismos si deberíamos actuar de la forma en la que instintivamente nos sentimos inclinados a actuar. Podemos preguntarnos: "Estoy inclinado a hacer A con el objetivo de lograr el fin F. Pero, ¿debería hacerlo?"

Esta habilidad – entender los principios en función de los cuales nos mostramos inclinados a actuar – es, de acuerdo con Korsgaard, parte de la esencia de la moralidad tal y como Kant la concibe:

[L]a capacidad de auto-gobierno normativo y el nivel más profundo de control intencional que dicha capacidad implica, es probablemente exclusiva de los seres humanos. Y es en el uso apropiado de esta capacidad – la habilidad de formar juicios sobre lo que deberíamos hacer y actuar en consecuencia – donde se encuentra la esencia de la moralidad, no en el altruismo o en la persecución del mayor bien.

Cualquier comportamiento que no esté sujeto a este tipo de auto-gobierno normativo no puede ser un comportamiento moral. Si una criatura es incapaz de reflexionar sobre lo que hace – de preguntarse a sí misma si, en unas circunstancias dadas, un acto es moralmente bueno, o de preguntarse a sí misma si lo que siente en esas circunstancias es el sentimiento moralmente apropiado para las mismas –, entonces no puede ser una criatura moral.

Gracia, la elefanta, presumiblemente es incapaz de reflexionar sobre sus motivaciones de esta manera. No puede reflexionar sobre sus motivaciones: no puede preguntarse a sí misma si éstas son buenas o malas, si debería aprobarlas o rechazarlas. Tampoco puede reflexionar en torno a la *conexión* entre sus motivaciones y sus acciones subsecuentes: si lo que se siente inclinada a hacer constituiría una respuesta adecuada a las motivaciones que la inclinan de ese modo. Desde la perspectiva Kantiana, las motivaciones de Gracia no pueden ser calificadas como motivaciones morales. Como mucho, sus motivaciones cuentan solamente como simulacros no morales de preocupación – sentimientos que se asemejan a las varias formas de preocupación en su perfil fenomenológico, pero que carecen del núcleo moral esencial que convertiría esta fenomenología en emociones propiamente morales. Sus motivaciones simplemente empujan a Gracia en esta u otra dirección – ella es el equivalente moral de un corcho balanceándose en los océanos de la motivación. Esas motivaciones solamente ejercerían presión *causal* en sus acciones resultantes; no ejercerían presión *normativa* alguna.

La “forma específica de auto-conciencia” que consiste en la comprensión de los principios sobre los cuales ella estaría inclinada a actuar lo cambiaría (supuestamente) todo. Gracia se alzaría sobre el océano, transformada de un individuo que se veía empujado en esta u otra dirección por sus motivaciones a uno capaz de alzarse sobre dichas motivaciones – observando y decidiendo serenamente sobre ellas.

Aristóteles ofrece una imagen similar, por lo menos en una resumida caracterización general. En el siguiente pasaje de la *Ética a Nicómaco*, Aristóteles enfatiza la complejidad psicológica de las virtudes:

Mientras que aquello que se realiza conforme a las virtudes no se realiza con justicia o con templanza porque estas mismas acciones sean de una determinada manera, sino porque, además, las realiza el agente estando de una determinada manera: en primer lugar si actúa con conocimiento; después, si lo hace por elección y si las elige por ellas mismas; y en tercer lugar, si también las realiza manteniéndose firme e incommovible.

Para que una acción sea la expresión de una virtud, no debe simplemente ser un ejemplo de lo que comúnmente sería considerado como una acción virtuosa (tener las "cualidades apropiadas"). Adicionalmente, el agente tiene que (a) saber que está realizando una acción virtuosa, y (b) realizar dicha acción porque es una acción virtuosa ("elegirlas por ellas mismas"), y (c) esta decisión tiene que ser la expresión de una disposición estable por parte del agente.

Es en la conjunción de las condiciones (a) y (b) donde descansa la hostilidad hacia la idea de que, al intentar ayudar a Eleonora, Gracia está actuando virtuosamente. En conjunción, de forma no muy diferente a como vimos en Kant, ambas condiciones imponen la reflexión como condición mínima del agente virtuoso. Para satisfacer estas condiciones, el agente tiene que comprender qué es una virtud, y estar motivado por esta comprensión a la hora de actuar, si es que la acción debe ser comprendida como la expresión de dicha virtud. Llamaré a esto – la conjunción de (a) y (b) – la *condición de reflexión* sobre la posesión/ expresión de una virtud:

Para que una acción Φ , realizada por un agente A, sea la expresión de una virtud, V, es necesario que A (i) sea capaz de comprender que Φ es un ejemplo de V, y (ii) A tiene que realizar Φ porque comprende que Φ es un ejemplo de V y desea ser virtuoso.

Aristóteles vincula estrechamente la condición de reflexión con el *asidero normativo* de una virtud. Según el punto de vista aristotélico, para que el intento de Gracia de ayudar a Eleonora cuente como un ejemplo de comportamiento virtuoso, debe ser cierto que (i) Gracia poseía el sentimiento motivacional correcto en dichas circunstancias, y (ii) el comportamiento resultante fue la respuesta moralmente apropiada a

esa motivación. La respuesta afirmativa a (i) y (ii) requiere la existencia de estándares de corrección que gobiernen tanto la posesión del sentimiento como la relación entre el sentimiento y la acción.

Para Aristóteles, el segundo elemento de normatividad se encuentra explicado por el hecho de que, cuando se dan como elementos constituyentes de una virtud, sentimiento y acción están relacionados internamente, no externamente. En otras palabras, el carácter normativo de la relación entre sentimiento y acción se explica por la forma de la estructura normativa de la virtud en sí misma. Así, si el primer elemento de normatividad – la idea de que el sentimiento debe ser apropiado para las circunstancias – puede ser explicado, el segundo elemento viene de suyo – como consecuencia de la estructura normativa de la virtud.

Aristóteles presenta dos explicaciones, las cuales se refuerzan mutuamente, del primer elemento de normatividad. Primero, enfatiza que las virtudes deben ser adquiridas, y esto puede hacerse correcta o incorrectamente. Esta es la razón por la que podemos ser alabados por la posesión de una virtud, y reprendidos por la ausencia de ella. Si las virtudes fueran como los sentidos, cuya posesión es una cuestión natural que no tiene nada que ver con nuestros esfuerzos, esas alabanzas o reprobaciones no tendrían ningún sentido.

Segundo, y más importante para nuestros propósitos, una virtud puede ejercerse mejor o peor. En un conocido pasaje de la *Ética a Nicómaco*, Aristóteles comenta:

También es propio de cualquiera y fácil el encolerizarse y el dar dinero y gastarlo, pero con quién, y en qué medida, y cuándo, y para qué, y cómo, ya no es propio de cualquiera ni tampoco fácil; por lo cual el bien es escaso, elogiabile y bello.

Si una acción cuenta o no como expresión de una virtud es un asunto sensible al contexto: solamente si la acción se realiza hacia la persona adecuada, en la medida correcta, en el momento adecuado, etc., contará como la expresión de una virtud. Pero éstos son a menudo asuntos difíciles que exigen un juicio. Este juicio puede realizarse correcta o incorrectamente; los errores siempre son posibles. Y es la posibilidad

de error en la aplicación de la virtud lo que hay debajo del estatus normativo de una virtud.

Por lo tanto, en relación con la cuestión de si el comportamiento de Gracia es virtuoso, Aristóteles presumiblemente nos recomendaría hacernos las siguientes preguntas: (1) ¿Sabe Gracia que lo que está haciendo es virtuoso? (2) ¿Se comporta Gracia de la forma en que lo hace porque desea ser virtuosa? (3) ¿Considera Gracia que Eleonora es la destinataria apropiada de este tipo de comportamiento? (4) ¿Piensa Gracia que su comportamiento es proporcional al sufrimiento de Eleonora? (Etcétera, etcétera...). Si Gracia es incapaz de albergar estos pensamientos (y otros relacionados), entonces el comportamiento de Gracia no cuenta como expresión de una virtud moral.

Tanto Kant como Aristóteles estarían de acuerdo en que los animales no cuentan como sujetos morales – sujetos de motivación moral – porque, y en la medida en que, son incapaces de examinar críticamente sus motivaciones y la relación entre sus motivaciones y sus acciones. Por debajo de este énfasis en la reflexión crítica se encuentran las ideas íntimamente relacionadas de normatividad y control. Sin la habilidad de reflexionar sobre sus motivaciones, no puede dársele ningún sentido a la idea de que la motivación es algo que un animal *debería* tener más que algo que simplemente *tiene*. Sin esta habilidad, las motivaciones de los animales simplemente les empujan primero en una dirección y luego en otra: ejercen una presión simplemente causal, pero no normativa, sobre su comportamiento. En ausencia de la capacidad para el razonamiento moral crítico, las motivaciones de los animales no tienen estatus normativo alguno; y la conexión entre estas motivaciones y su comportamiento es, más que normativo, simplemente causal.

Aristóteles y Kant son dos ejemplos prominentes del punto de vista ortodoxo en torno a la motivación moral. En lo que resta de este artículo intentaré, como le gusta decir a Bill Lycan, moler este punto de vista ortodoxo hasta convertirlo en polvo fino.

6. El idiota

Consideremos a una persona que, en consonancia con el príncipe de *El Idiota* de Dostoievski, voy a llamar 'Mishkin'. Prima facie, Mishkin posee el alma de un príncipe: a lo largo de su vida, realiza muchos actos que parecen amables y compasivos. Él realiza estos actos porque es el sujeto de sentimientos o emociones que – nuevamente, como mínimo *prima facie* – parecen ser amables y compasivos. Cuando ve a alguien sufriendo, se siente impelido a actuar para terminar o aliviar dicho sufrimiento. Cuando observa a alguien feliz, él mismo se siente feliz por lo que ve. Si puede ayudar a que alguien consiga lo que persigue sin dañar a nadie, le ayudará porque se da cuenta de que disfruta haciéndolo. En resumen, Mishkin deplora el sufrimiento y se regocija en la felicidad ajena. Sus acciones reflejan, y están causadas por, estos sentimientos. Lo que Mishkin no hace, sin embargo, es someter sus sentimientos y acciones a escrutinio crítico moral alguno. De esta forma, él jamás se plantea algo así como: "¿Es lo que estoy sintiendo el sentimiento adecuado para la situación actual – es decir, es lo que estoy sintiendo igual a lo que debería estar sintiendo?" Tampoco se plantea cuestiones como: "¿Es lo que me propongo hacer en estas circunstancias la acción (moralmente) correcta (teniendo en cuenta todos los factores)?" Él no se plantea estas cuestiones, supongámoslo, porque es incapaz de planteárselas. Sus relaciones con los otros operan a un nivel más visceral. Lo que podríamos pensar del perfil moral de Mishkin sería aproximadamente algo como esto:

(M1) (i) Mishkin realiza acciones que parecen buenas, y (ii) la motivación de Mishkin para realizar estas acciones consiste en sentimientos o emociones que parecen ser buenos, pero (iii) Mishkin es incapaz de someter a escrutinio moral crítico tanto sus acciones como sus sentimientos.

¿Son suficientes (i)-(iii) para calificar a Mishkin como un sujeto moral – es decir, para considerar sus motivaciones como motivaciones morales? Si hemos sido persuadidos por Aristóteles y/o Kant, nuestra respuesta sería negativa. Pero, ¿cómo de razonable es este rechazo? Mishkin se

pasa la vida ayudando a los otros, y lo hace porque se deleita con la felicidad de los otros y deplora su sufrimiento. Necesitamos distinguir (a) si alguien puede ser clasificado como sujeto moral, en el sentido de encontrarse motivado por preocupaciones morales, y (b) si alguien puede ser clasificado como sujeto moral precisamente en la forma en que lo hacemos nosotros. ¿Es posible que Aristóteles y Kant describieran solamente una forma de ser clasificado como sujeto moral? ¿Es posible que haya otras?

Estas cuestiones, no obstante, nos invitan simplemente a expandir nuestra concepción de lo que es un sujeto moral. Probablemente tanto Aristóteles como Kant hubieran respondido con un firme, "¡No gracias!". Para que pueda incomodarles, la invitación necesita ser reforzada y convertida en algo parecido a una oferta que no puede ser rechazada.

Como comienzo de este proceso de refuerzo, merece la pena señalar que existe un sentido claro en el que (M1) le hace un flaco favor a Mishkin. Esto tiene que ver con la caracterización de sus acciones y sentimientos como acciones y sentimientos que *parecen* ser buenos. La motivación para una caracterización de este tipo está muy clara. Aquellos que rechazan la idea de que Mishkin es un sujeto moral rechazarán también la idea de que sus emociones y acciones cuenten como buenas (por ejemplo, como amables o compasivas) – una descripción así solamente sería cierta para los sujetos morales. Sin embargo, supongamos que alguien – Marlow, en consonancia con el escrutinador de motivaciones que narra algunas de las novelas de Joseph Conrad – es capaz del tipo de reflexión moral crítica de la que Mishkin es incapaz. Y supongamos que, en base a esa reflexión suya, Marlow fuera a refrendar el mismo tipo de sentimientos y acciones que Mishkin. Es decir, supongamos que para cualquier circunstancia dada C, Mishkin posee el sentimiento S y, como resultado, realiza la acción A. Marlow, un hábil escrutinador de sus sentimientos y acciones, llega independientemente a la conclusión de que en la circunstancia C, es moralmente correcto poseer el sentimiento S y realizar la acción A. Supongamos, además, que Marlow – al cual podríamos considerar como un espectador *ideal* o *racional* – invariablemente alcanza la conclusión moral correcta (o, si lo prefieren, la alcanza la mayoría de las veces).

En otras palabras, Mishkin posee la disposición estable de experimentar ciertas emociones y, consecuentemente, realizar ciertas acciones en una gama de circunstancias identificables, y Marlow – nuestro espectador racional ideal – llega independientemente a la conclusión de que, para todas (o la mayoría) de circunstancias identificadas, poseer esos sentimientos y realizar esas acciones es moralmente correcto. Si este fuera el caso, podríamos reforzar (M1) como sigue:

(M2) (i) Mishkin realiza acciones que son, de hecho, buenas, y (ii) la motivación de Mishkin para realizar dichas acciones consiste en emociones o sentimientos que son, de hecho, los adecuados en esas circunstancias, pero (iii) Mishkin es incapaz de someter a escrutinio moral crítico alguno estos sentimientos y acciones.

La transición de (M1) a (M2) es significativa: quiere decir que Mishkin tiene razones *externas* para sus acciones. Una razón interna para una acción es aquella que promueve ciertos motivos del agente – si el agente actualmente posee dichos motivos, o si llegaría a poseer dichos motivos siguiendo ‘una ruta deliberativa sensata’. Por otro lado, un agente tendría una razón externa para Φ si (i) tuviera alguna razón para Φ , y (ii) ninguno de sus motivos o intereses fueran promovidos por Φ .

En el caso de (M1), puede argumentarse que Mishkin no tiene razones morales para realizar sus acciones: El Mishkin de (M1) está a merced de sus motivaciones; no tiene control alguno sobre su aprobación o resistencia a las mismas, y así ellas ejercen solamente presión causal, pero no normativa, sobre sus acciones. Pero la normatividad es esencial para que algo sea una razón más que una causa. Las motivaciones de Mishkin, por lo tanto, aunque pueden ser la causa de su comportamiento, no son razones morales. Sin embargo, (M2) introduce un nuevo tipo de razón moral: el Mishkin de (M2) tiene razones morales externas para realizar sus acciones. Mishkin no puede contemplar esas razones y, por tanto, no pueden ser razones internas. No obstante, son razones morales que existen para que Mishkin haga lo que hace.

De modo que, ahora, podemos repetir la pregunta: ¿afirmaríamos que Mishkin – el caracterizado por (M2) – no es un sujeto moral? Él realiza acciones que son, de hecho, moralmente buenas, y las realiza partiendo de una motivación que es, de hecho, moralmente encomiable. Hay razones – razones morales – que conectan sus sentimientos y acciones: lo que sucede es que son razones externas. ¿Deberíamos entonces mostrarnos tan confiados en nuestra afirmación de que Mishkin no cuenta como un ser moral? Esto constituye una invitación a pensar la moralidad de una manera determinada; ¿pero puede esta invitación ser razonablemente declinada?

Tal vez sí. Pero un filósofo moral escéptico que todavía desee declinar esta invitación se encuentra comprometido con la idea de que, para ser un sujeto moral, es condición necesaria poseer razones internas a la hora de actuar. Estas razones internas están disponibles para el escrutinio racional; son razones que uno no solamente puede refrendar, sino que uno tendría, en cierta medida, que haber refrendado de hecho. Esta idea está, por supuesto, lejos de ser irrazonable. La existencia de razones externas es controvertida. Bernard Williams, que las introdujo en el discurso filosófico, argumentó repetidamente que tales cosas no existen.

Es más, incluso alguien que aceptara la existencia de razones morales externas podría mostrarse reacio a aceptar la afirmación de que Mishkin es un sujeto moral. Existe, uno podría argumentar, cierto déficit que acompaña la carencia de Mishkin de razones internas. Hay un sentido claro en el cual, aunque Mishkin capta “correctamente” las cosas – es decir, en la circunstancia dada C, realiza las acciones moralmente correctas y experimenta las emociones moralmente adecuadas – las capta correctamente por *accidente*. Mishkin actúa en base a las emociones que posee, y no ha escudriñado ni las emociones, ni las acciones, ni la conexión entre las emociones y las acciones. Así, mientras que hay un sentido en el cual Mishkin capta las cosas correctamente, moralmente hablando, no capta las cosas correctamente *por las razones adecuadas*. Y eso es lo que se requiere para que Mishkin pueda ser considerado un sujeto moral – un sujeto de motivación moral como es debido.

La imagen de un sujeto moral reflexivo como alguien que puede captar las cosas correctamente – por ejemplo, identificar la conclusión moral correcta – *por las razones adecuadas* es una imagen que muchos encontrarán persuasiva. Es más, esta es una forma de desarrollar la conexión entre normatividad y control que, he argumentado, era central en los enfoques ortodoxos sobre la motivación moral. Así que, tratemos de acabar con el elemento contingente ofensivo en las decisiones morales de Mishkin. Hasta cierto punto, podemos hacerlo mediante la internalización de las razones de Mishkin – aunque, crucialmente, no hasta el punto de convertirle por completo en un sujeto moral reflexivo.

Supongamos que la manera en que Mishkin “capta correctamente las cosas” – es decir, poseyendo los sentimientos correctos y realizando las acciones adecuadas en unas circunstancias dadas – no sea tan accidental como pueda parecer. Supongamos que Mishkin tenga buenas razones para sus sentimientos e igualmente buenas razones para las acciones resultantes de los mismos. Lo que sucede es que esas razones no están disponibles para su propia conciencia o escrutinio racional. Son razones que han sido encarnadas en el proceso de operaciones inferencial a nivel infra-personal y no-consciente de Mishkin. Por ejemplo, si nuestros gustos se inclinaran hacia lo modular (aunque es completamente opcional pensar de esta manera), podríamos suponer que Mishkin posee un “módulo moral”, cuyas operaciones son cognitivamente impenetrables – es decir, que funcionan al margen de las posteriores operaciones que tienen que ver con la formación de creencias y conceptos. Aun así, las operaciones del módulo son confiables y llevan a conclusiones morales correctas con el mismo grado de éxito de Marlow, nuestro espectador racional/ideal. En esas circunstancias, podríamos enmendar (M2) de esta manera:

(M3) (i) Mishkin realiza acciones que son buenas, y (ii) la motivación de Mishkin para realizar esas acciones consiste en emociones o sentimientos que son moralmente adecuados para esas circunstancias, y (iii) Mishkin tiene razones para poseer esos sentimientos y realizar esas acciones en dichas circunstancias, pero (iv) Mishkin no es consciente de esas razones.

(M2) otorgó a Mishkin razones externas para sus acciones. (M3), en sentido claro, internaliza esas razones. De ahora en adelante ya no son simplemente razones que existen para Mishkin. Son razones que Mishkin *posee*. Las posee no en el sentido de que se encuentran disponibles para su escrutinio crítico moral, sino en el sentido de que se encuentran encarnadas en las operaciones de procesamiento, infrapersonales e inconscientes, que ocurren en su módulo moral. Estas operaciones culminan en sus, a nivel personal, sentimientos y acciones morales conscientes.

De modo que, una vez más, nos enfrentamos a la pregunta: ¿Queríamos rechazar que el Mishkin caracterizado por (M3) es un sujeto moral? ¿No resultaría más razonable suponer que este Mishkin es un sujeto moral simplemente algo diferente de los críticos evaluadores morales por los que nos tenemos a nosotros mismos? Mi intuición me dice que, llegados a este punto, no considerar a Mishkin como sujeto moral se ha vuelto irrazonable. Mishkin es un sujeto moral; sólo que algo diferente de aquel que nosotros – o como mínimo Kant y Aristóteles – consideramos que somos.

Sin embargo, esto es solamente una intuición. La objeción principal a la misma probablemente se derive del pensamiento de que mientras (M3) podría haber eliminado el elemento contingente ofensivo del perfil moral de Mishkin, no lo ha hecho de la manera adecuada. (M3) podría mitigar la contingencia, pero no hace nada por paliar la ausencia de control de Mishkin sobre lo que siente y la manera en que actúa. Mishkin no tiene ni idea de qué es lo que sucede en su módulo moral. No tiene autoridad alguna sobre su funcionamiento, y por tanto no tiene ningún control sobre sus producciones. Pero el control, he argumentado, es una condición necesaria de la normatividad. Si Mishkin no tiene ningún control sobre el funcionamiento de su módulo moral, entonces los contenidos de este módulo no tienen ningún estatus normativo: no son el tipo de cosas que Mishkin debería promover o a las que debería resistirse. En última instancia, el rechazo de que las motivaciones de Mishkin cuenten como motivaciones morales encuentra su fundamento en las ideas afines de autoridad y control.

Muy bien: comencemos a ejercer presión sobre las ideas de autoridad y control.

7. Super-visión ciega moral³

Tratemos de identificar de forma más precisa la naturaleza de la diferencia entre el Mishkin de (M3) y un sujeto moral reflexivo. Un sujeto moral reflexivo – representado por Marlow – se caracterizaría, presumiblemente, por (M4):

(M4) (i) Marlow realiza acciones que son buenas, y (ii) la motivación de Marlow para realizar dichas acciones consiste en emociones o sentimientos que son moralmente apropiados en esas circunstancias, y (iii) Marlow tiene buenas razones para poseer esos sentimientos y realizar esas acciones en dichas circunstancias, y (iv) Marlow es consciente de esas razones.

La diferencia entre Mishkin y Marlow entonces pertenece solamente a la condición (iv). Si el control sobre las motivaciones es lo que separa a Mishkin de un sujeto moral genuino, entonces es la condición (iv) la que tiene que hacer posible este control. Al caracterizar esta diferencia entre Mishkin y Marlow, podríamos mantenernos en la jerga de la modularidad: las operaciones del módulo moral de Marlow se encuentran disponibles para subsecuentes operaciones (conscientes) de formación de creencias, en tanto que las del módulo de Mishkin no lo están. El módulo moral de Marlow es *cognitivamente penetrable*, mientras que el de Mishkin no lo es.

Hay dos formas mediante las cuales podríamos explicar la impenetrabilidad cognitiva del módulo de Mishkin.

(1) Las operaciones que ocurren en el módulo moral de Mishkin no son *fenoménicamente conscientes*. Es decir, *no existe algo así* como ser el sujeto de estas operaciones.

Por el contrario, para Marlow, existe algo así como dedicarse a la deliberación moral. Esto, de acuerdo con la primera sugerencia, constituye la diferencia que la condición (iv) trata de captar.

La segunda forma de explicar la antedicha carencia de disponibilidad se formula en términos de la incapacidad de Mishkin a la hora de formar *pensamientos de orden superior* en torno a sus deliberaciones morales:

- (2) Cuando Mishkin se enfrasca en deliberaciones morales, no tiene pensamiento de orden superior alguno que le indique que eso es lo que está haciendo.

Esta sección estará dedicada a (1). Mishkin tiene cierto tipo de *acceso* a las operaciones de su módulo moral. Como quedó caracterizado en (M3), estas operaciones son responsables de los sentimientos de Mishkin, y conectan esos sentimientos con sus acciones subsecuentes al proveerle de razones – inaccesibles para Mishkin – para realizar dichas acciones. En otras palabras, las operaciones de su módulo moral determinan su toma de decisiones real/actual. Esto, en la terminología asociada a Block, es suficiente para mostrar que estas operaciones son conscientemente accesibles: donde, por decirlo toscamente, un estado o proceso es conscientemente accesible si está jugando un rol activo en los procesos psicológicos reales/actuales y las acciones resultantes. Sin embargo, estas operaciones no son fenoménicamente conscientes. No hay nada para Mishkin que sea algo así como experimentar las operaciones que ocurren en su módulo moral. Marlow tiene también acceso consciente a las operaciones de su módulo. Pero, adicionalmente, existe para Marlow algo así como la experimentación de dichas operaciones: sus operaciones son operaciones fenoménicamente conscientes.

Si esta es la diferencia entre Mishkin y Marlow, entonces nuestra pregunta es clara y concisa: ¿la consciencia fenoménica (o la ausencia de la misma) de los procesos morales de toma de decisiones nos proporciona una vía legítima para establecer quiénes son sujetos morales y quiénes no?

Las posibilidades iniciales de una respuesta afirmativa son, sospecho, poco promisorias. He argumentado que es la idea de autoridad y control la que subyace en el centro de la demarcación. El Mishkin de (M3) no es un sujeto moral en la medida en que carece de autoridad y control sobre sus procesos morales de toma de decisiones. Pero resulta implausible suponer que este tipo de fenomenología moral deliberativa – poseída por Marlow pero no por Mishkin – producirá, por sí misma, el control sobre las motivaciones de uno. Por ejemplo, el determinista radical no necesita rechazar la fenomenología de la libertad. Simplemente rechaza

que nuestra sensación de ser libres añada nada a nuestra libertad real. De forma similar, la fenomenología de la deliberación moral puede asegurar nuestra *sensación* de que tenemos control sobre nuestras motivaciones, pero no la afirmación más fuerte de que efectivamente *tenemos* control sobre las mismas. Si es el control sobre las motivaciones lo que separa a Marlow de Mishkin, entonces esto no lo encontraremos en la fenomenología moral.

Además, la idea de que la fenomenología de la deliberación moral sirva para delimitar los sujetos morales de los no morales es implausible por razones independientes – aunque no respaldemos la conexión entre normatividad y control. Éste es, pienso, un camino útil a la hora de pensar sobre el Mishkin de (M3). Mishkin está caracterizado por una forma de *super-visión ciega moral*. En su sentido ordinario, “visión ciega” se refiere al déficit de consciencia fenoménica. Debido a lesiones en el córtex visual, un sujeto con visión ciega afirmará no tener consciencia alguna de la existencia de un objeto en cierto lugar de su campo visual. Sin embargo, cuando se les pregunta dónde se encuentra el objeto en cuestión, los sujetos lo adivinan en un porcentaje superior al esperable en una elección aleatoria. Los sujetos con visión ciega no son funcionalmente equivalentes a aquellos que tienen una visión normal. No lo son, habitualmente tendrán grandes dificultades para moverse por el entorno, por lo menos en condiciones normales. Aun así, la visión ciega suele considerarse típicamente como indicativa de la posibilidad de que la consciencia fenoménica y el tipo de habilidades funcionales que constituyen el acceso a la consciencia puedan ir separadas.

Basándonos en esto, podemos imaginar un caso de lo que podemos llamar super-visión ciega. Una persona con super-visión ciega no tiene consciencia fenoménica visual alguna. Si le preguntaras qué es lo que ve, te contestaría: nada. No obstante, se mueve fácilmente por el entorno, evitando los objetos que se cruzan en su camino, etc. Si le preguntas al sujeto: ¿cómo sabías que ese objeto – el objeto que acaba de esquivar – estaba ahí? Respondería: simplemente tenía el presentimiento de que había algo ahí. Lanzas una pelota de béisbol hacia él, y la atrapa fácilmente. Le preguntas: ¿Cómo pudiste hacer eso? Y responde:

simplemente tenía el presentimiento de que había algo ahí. Así es como la persona con super-visión ciega consigue desenvolverse en su entorno: basándose en sentimientos (intencionales). Pero estos sentimientos son completamente fidedignos – o por lo menos tan fidedignos como las impresiones visuales de las personas videntes. Como resultado, el sujeto con super-visión ciega es capaz de moverse por el entorno con un nivel de eficacia más o menos igual al de la persona vidente.

¿Puede ver el sujeto con super-visión ciega? Esta pregunta es ambigua. Por una parte, puede significar: ¿tiene el sujeto con super-visión ciega consciencia visual? La respuesta negativa a esta cuestión es plausible. Sin embargo, la pregunta puede significar también: ¿es alguien con super-visión ciega un sujeto visual? Responder negativamente a esta cuestión resulta mucho menos plausible. De este modo, al preguntarle al sujeto con super-visión ciega qué objeto está delante de él, éste dirigiría sus ojos en la dirección apropiada. La luz podría llegar hasta sus ojos, y un mensaje podría ser enviado al córtex visual a través del nervio óptico. Varias operaciones de procesamiento de información ocurrirían ahí, y el resultado sería un sentimiento con, por ejemplo, el siguiente contenido: “¡Es una col!”. Debemos hacer notar también que los procesos en cuestión son sensibles solamente a las propiedades visuales destacadas de los objetos. No es como si, por ejemplo, los procesos le dieran acceso a las entrañas de la col, sino solamente a las propiedades que podrían ser vistas por alguien con capacidad visual normal.

Resultaría implausible decir que, en este caso, el sujeto con super-visión ciega no está viendo. No es que simplemente detecte objetos, sino que la manera en que los detecta resulta decisiva. Así, por ejemplo, no localiza los objetos mediante ecolocalización. Utiliza sus ojos, y los mensajes de sus ojos son procesados en el córtex visual de su cerebro. Simplemente sucede que estos procesos que habitualmente dan lugar a una percepción fenomenológica visual, en esta ocasión dan lugar a un sentimiento de, supongamos, un carácter más visceral. Antes que negar que el sujeto está viendo, considero mucho más plausible decir que esto es precisamente en lo que ver consiste para el sujeto con super-visión ciega.

(M3), en efecto, ofrece la caracterización de alguien que sufre de lo que podríamos llamar *super-visión ciega moral*. Mishkin realiza las acciones adecuadas (a saber, moralmente correctas), siente los sentimientos correctos, e incluso tiene las razones adecuadas para realizar esas acciones y sentir esos sentimientos, pero no es consciente de esas razones, y por tanto no tiene ni idea de por qué se siente de la forma en que lo hace, o por qué debería hacer las cosas que hace. Moralmente hablando, sin embargo, continúa captando correctamente las cosas. La pregunta es: ¿querríamos decir que Mishkin no es un sujeto moral – que no está ni puede estar motivado por consideraciones morales?

Mishkin, como sujeto con *super-visión ciega moral*, parece ser un sujeto moral en la manera en que una persona con *super-visión ciega* es un sujeto visual. La capacidad visual es compatible con diferentes fenomenologías visuales. De forma parecida, deberíamos aceptar que la capacidad de ser moral – de estar motivado por consideraciones morales – es compatible con fenomenologías morales diferentes. Antes que rechazar que la persona con *super-visión ciega moral* sea un sujeto moral, resulta mucho más plausible afirmar que, para ellos, ésta es precisamente su forma de ser un sujeto moral. Ser un sujeto moral no está tan estrechamente ligado a la fenomenología como para que podamos denegar a alguien el estatuto de sujeto moral por el simple hecho de que su fenomenología moral asociada difiera de la nuestra.

8. Sujetos morales y meta-cognición

Hay otra diferencia entre Marlow y Mishkin; una que podría, *prima facie*, parecer más significativa. Marlow tiene, mientras que Mishkin carece de, *pensamientos de orden superior* sobre sus motivaciones y procesos de toma de decisiones. Es decir, Marlow es capaz de albergar pensamientos del siguiente tipo: “Estoy siendo motivado a realizar Φ por la razón R”. Para tener este pensamiento, Marlow tiene que ser capaz de albergar otros pensamientos – por ejemplo, “estoy motivado a realizar Φ ” y “soy el sujeto de la razón R” (o, de forma equivalente: “La razón R es una

razón que yo refrendo”). Estos son pensamientos de orden superior – pensamientos sobre otros estados mentales (en este caso concreto motivaciones y razones). En este sentido Mishkin es incapaz de albergar pensamientos de orden superior.

¿Son los pensamientos de orden superior sobre una motivación, M, el tipo de cosas que pueden convertir M en una motivación propiamente moral? En este contexto, la noción de ‘conversión’ puede ser entendida de dos maneras: ¿Es el pensamiento de orden superior sobre la motivación (i) *necesario* o (ii) *suficiente* para que esa motivación cuente como una motivación propiamente moral? En el argumento que presentaré a continuación, me centraré en la condición de suficiencia, y argumentaré que los pensamientos de orden superior no son suficientes para que las motivaciones se transformen en motivaciones morales. No obstante, el fracaso de la suficiencia tiene implicaciones más amplias. Las razones para el fracaso de la suficiencia, tal y como trataré de mostrar, indican que los pensamientos de orden superior *no son simplemente el tipo de cosas adecuadas* a la hora de constituir las motivaciones como motivaciones morales. Por lo tanto, las razones para el fracaso de la suficiencia sugieren fuertemente un fallo también en la necesidad.

Recapitemos la dialéctica de la cuestión. En ausencia de las habilidades meta-cognitivas relevantes, Mishkin se encuentra “a merced” de sus motivaciones. Ellas le empujan en una u otra dirección, y él no es consciente del rol que juegan y así no tiene control alguno sobre dicho rol. Sin embargo, las habilidades meta-cognitivas – la habilidad de albergar pensamientos de orden superior sobre sus motivaciones – transforman a Marlow. Él es capaz de alzarse sobre las refriegas motivacionales, observando, juzgando y evaluando sus motivaciones y decidiendo la medida en la que les va a permitir determinar sus decisiones y acciones. Marlow puede preguntarse a sí mismo si debería seguir los dictados de una motivación dada, si debería abrazarla o resistirse a ella. Mishkin es incapaz de hacer esto y, por tanto, no se le puede dar sentido alguno a la idea de que *debería* hacerlo. Deber implica poder. La habilidad de Marlow de someter sus motivaciones a un escrutinio de este tipo, en consecuencia, le otorga a sus motivaciones un estatus normativo del que las de Mishkin carecen.

Sin embargo, ¿qué sucedería si las habilidades meta-cognitivas de Marlow no fueran del todo como se sugiere en este enfoque? ¿Qué pasaría si fuesen, en ciertos aspectos cruciales, como las motivaciones que ellas toman como sus objetos? Marlow puede indudablemente pensar para sí mismo pensamientos como, “¿Qué es lo que me está motivando a hacer esto?” y “¿Es esta motivación algo en base a lo que debería actuar o es algo que debería rechazar?” No obstante, supongamos que las respuestas que da a estas cuestiones son respuestas sobre las que él tiene poco control: al ser, por ejemplo, crucialmente sensible a las características de la situación en la que se encuentra – características sobre las cuales tiene solamente una conciencia débil y cuya influencia sobre él escapa ampliamente a sus capacidades cognoscitivas conscientes. En otras palabras, mientras Marlow podría todavía observar y evaluar sus motivaciones, esta observación y evaluación se encontraría tanto oscurecida como determinada por factores contextuales de los cuales es muy poco consciente y sobre los que tiene muy poco control. En esas circunstancias, parecería que Marlow se encuentra “a merced” de sus evaluaciones meta-cognitivas en torno a sus motivaciones. Estas evaluaciones lo empujan en una u otra dirección – le hacen evaluar sus motivaciones de una u otra manera – pero trabajan de forma subterránea. Él es levemente consciente de, y tiene muy poco control sobre, sus funcionamientos.

Estoy, por supuesto, aludiendo a los influyentes trabajos recientes en psicología moral de orientación, en términos generales, *situacionista*. Los enfoques situacionistas acerca del sujeto moral pueden ser contrastados con los enfoques *disposicionalistas*. De acuerdo con estos últimos, muy resumidamente, el sujeto moral está constituido, en parte, por un conjunto de disposiciones que son internas al sujeto y el sujeto es responsable de su formación y preservación. Estas disposiciones componen lo que podríamos llamar el *carácter* de la persona. Incluidas en estas disposiciones estarán algunas concernientes a la evaluación de las motivaciones. Así que, el carácter del sujeto moral está, en parte, compuesto de disposiciones que clasifican las motivaciones dadas como motivaciones en base a las cuales deberíamos actuar o motivaciones que deberían ser rechazadas.

Los enfoques situacionistas, sin embargo, ven las cosas de forma diferente. De acuerdo con la (probablemente irrealista) versión fuerte del situacionismo, el carácter de una persona es tan maleable como las situaciones en las que se encuentra a sí misma. Cambiemos la situación y las disposiciones del sujeto a evaluar moralmente las motivaciones de una forma antes que de otra cambiarán también. Así, podríamos imaginar a Marlow como un participante en el Experimento de la Cárcel de Stanford o como guardia en Abu Ghraib. Si el enfoque situacionista fuera correcto, deberíamos esperar que las evaluaciones de Marlow sobre sus motivaciones variaran junto con estas variaciones en las circunstancias.

Así pues, la idea de que las habilidades meta-cognitivas de Marlow le confieren control y, por tanto, normatividad sobre sus motivaciones – normatividad de la que por tanto carecen las motivaciones de Mishkin – parece, como poco, rehén de la fortuna empírica. Si Marlow está, como argumenta el situacionista, a merced de sus evaluaciones meta-cognitivas sobre sus motivaciones de la misma forma en la que Mishkin ha sido retratado como a merced de sus motivaciones, entonces no está claro cómo la posesión de habilidades meta-cognitivas puede conferir control alguno sobre las motivaciones de Marlow; y, por tanto, está igualmente poco claro cómo pueden imbuir normatividad a estas motivaciones.

No obstante, por debajo de estas preocupaciones empíricas se encuentra un asunto conceptual más profundo e importante para nuestros propósitos. La apelación a la meta-cognición trata de explicar la normatividad de nuestras emociones por medio de nuestro control sobre las mismas. Esto pasa por alto el hecho de que el problema del control, surgido en el nivel de las motivaciones, va a repetirse también en el nivel de nuestras evaluaciones sobre esas motivaciones (segundo-orden). Esta incapacidad de ver lo que es, en última instancia, un punto bastante claro constituye un ejemplo clásico de la atracción que sobre nosotros ejerce el meta-nivel. El fallo es el resultado de suponer que algo mágico ocurre al movernos del primer-orden al meta-nivel – magia que inmuniza el meta-nivel ante las tribulaciones del primer-orden.

Se suponía que la cognición de orden-superior permitía a Marlow alzarse sobre la refriega motivacional y juzgar tranquilamente sus motivaciones, transformándolas así en objetos normativos. Sin embargo, no hay razón alguna para suponer que la cognición de orden-superior se alce sobre esta refriega motivacional. Si las motivaciones de primer-orden pueden empujar a Mishkin de esta u otra manera, entonces, las evaluaciones de segundo-orden en torno a esas motivaciones pueden hacer exactamente lo mismo con Marlow. Si Mishkin se encuentra "a merced" de sus motivaciones de primer-orden, entonces, lógicamente, Marlow está de forma similar "a merced" de sus evaluaciones de segundo-orden en torno a dichas motivaciones. En resumen, las evaluaciones de segundo-orden en torno a nuestras motivaciones de primer-orden no pueden alzarnos sobre la refriega motivacional que consideramos endémica en el ámbito de las motivaciones de primer-orden, por la simple razón de que podemos estar motivados a evaluar nuestras motivaciones de una forma antes que de otra. Nada mágico o milagroso ocurre en el meta-nivel: cualquier característica, patrón y, más importante para nuestros propósitos, deficiencia que encontremos en el nivel de primer-orden, será simplemente reiterada en el nivel de segundo-orden.

9. Razones y práctica

Hay una tercera diferencia hecha posible por la condición (iv): habilita a Marlow a entrar en la *práctica* de explicar tanto sus comportamientos morales como los de otros por medio de *razones*. Cuando Mishkin se presenta ante una situación en la cual un individuo está sufriendo, siente un profundo deseo de tratar de aliviar su sufrimiento. Pero Mishkin no puede explicar, ni para sí mismo ni para los demás, qué es lo que está haciendo mediante el ofrecimiento de razones para su comportamiento. No puede albergar pensamientos como: "Hice eso porque quería aliviar el sufrimiento de Smith". Marlow, por otra parte, sí puede hacerlo. Así pues, Marlow puede formar parte de lo que podríamos llamar *la práctica de ofrecer razones*.

Existe una tradición de pensamiento sobre razones para la acción (y las creencias), encarnada en pensadores como Wittgenstein y Sellars, que hace que formar parte de la práctica de ofrecer razones sea necesario para poseer razones para la acción (y las creencias). Un individuo que no pudiera tomar parte en una práctica de este tipo pertenecería, tal y como Sellars dice, al *espacio de las causas* pero no al *espacio de las razones*. Pero si un individuo así no tiene razones para lo que hace, *a fortiori* no tiene razones morales. Para poseer razones morales, es necesario formar parte de *la práctica de ofrecer razones*: la práctica de mencionar consideraciones morales y explicar el comportamiento mediante la apelación a conceptos normativos como lo correcto y lo incorrecto.

Definamos la *hipótesis de la práctica* tal y como sigue: la pertenencia a la práctica de ofrecer razones, *P*, (en lo sucesivo simplemente "práctica"), por un sujeto, *S*, es una condición necesaria para que los estados motivacionales de *S* cuenten como razones. O sea, una condición necesaria para que un estado motivacional cuente como razón es que éste sea poseído por alguien que forme parte de dicha práctica. Esta hipótesis está motivada por cierta forma de entender la conexión entre razones, contenidos y prácticas. Una explicación del comportamiento que cite razones, en oposición a causas, es aquella que se centra en el contenido de la razón: aproximadamente, la razón tiene un contenido que hace que el comportamiento sea inteligible. Este contenido se transfiere entonces a las descripciones que hacemos del comportamiento resultante. La apelación a la práctica es, tradicionalmente, un intento de explicar la posibilidad del contenido o el significado. Consideremos, por ejemplo, la apelación a la práctica que es la culminación de las cavilaciones de Wittgenstein sobre el seguimiento de reglas. Nos sentimos tentados a suponer, especialmente en ausencia de alternativa viable alguna, que el significado de una expresión es una función de las reglas que seguimos al emplearla. Pero, ¿en qué consiste seguir una regla? Habiendo rechazado – correctamente – las interpretaciones mentalistas de lo que es seguir una regla, Wittgenstein se encuentra impelido a apelar a la práctica. Seguir una regla para el uso de un signo es la práctica de ajustar el

uso que uno hace de ese signo al uso acostumbrado del mismo. Sin la práctica, no habría contenido. Sin contenido no habría razones para la acción, simplemente causas de comportamientos.

Sin embargo, si esta es la conexión prevista entre formar parte de una práctica y ser el sujeto de ciertas razones, la conexión es solamente tan segura como lo sea la apelación a la práctica. Y esto debe enfrentarse a dificultades abrumadoras. En particular, la apelación es circular: presupone el contenido más que explicarlo. La práctica es lo que hacemos. Lo que hacemos son acciones. Las acciones son individuadas por las intenciones (o voliciones, aspiraciones, complejos de creencias-deseos, etc.). Pero las intenciones son individuadas por su contenido. Una intención es distinta (como tipo) de otra, si y solamente si, en las cláusulas subordinadas de sus adscripciones figuran proposiciones diferentes. La apelación a la práctica, por tanto, no puede explicar cómo el contenido es posible por la simple razón de que el contenido se presupone. Pero si la apelación a la práctica no puede explicar el contenido, no tenemos razón alguna para suponer que el contenido no puede darse en ausencia de práctica alguna.

10. Condiciones de la subjetividad moral

He argumentado que la habilidad de escudriñar críticamente las motivaciones propias no es una condición necesaria para que un ser pueda ser considerado un sujeto moral. Esta habilidad no hace, incluso no puede hacer, lo que se supone que debería hacer: conferir control, y por tanto normatividad, a las motivaciones de uno. Si esto es correcto, toda una tradición de pensamiento sobre en qué consiste la posesión de razones morales para las acciones de uno tiene que ser rechazada. Y este rechazo allana el camino para que los animales puedan ser considerados sujetos morales.

Mishkin reúne los requisitos para poseer razones morales para sus acciones cuando satisface tres condiciones: (1) la posesión de una *sensibilidad* apropiada ante las características pertinentes de su entorno,

donde (2) esta sensibilidad es *normativa*, y (3) está cimentada en un *mecanismo confiable*. Primero, las emociones de Mishkin *rastrear* las características que hacen buena – o mala – una situación. Cuando Mishkin encuentra que alguien está sufriendo siente un profundo deseo de aliviar su sufrimiento. Cuando se encuentra con alguien feliz, se regocija en ello, y así sucesivamente. La noción de “rastrear” o “seguir la pista” puede aprehenderse por medio de contrafácticos. Por ejemplo, *ceteris paribus*, si la situación no contuviera ninguna persona que estuviese sufriendo, entonces Mishkin no tendría, específicamente, esta emoción y el deseo resultante, y así sucesivamente.

Segundo, la sensibilidad de Mishkin es normativa. Marlow, nuestro espectador imaginario ideal, refrenda las respuestas emocionales de Mishkin ante las circunstancias del entorno, y lo hace porque está, tanto como cualquiera pueda estarlo, en posesión de una teoría moral correcta.

Tercero, la sensibilidad normativa de Mishkin no ocurre por accidente. Esta sensibilidad está cimentada en un mecanismo – un módulo moral si se prefiere – que genera ciertas emociones en unas circunstancias dadas.

Si la condición (4) – la condición meta-cognitiva que era satisfecha por Marlow pero no por Mishkin – falla, entonces sospecho que en el centro de la idea de la motivación moral nos encontramos con *la sensibilidad normativa cimentada en un mecanismo confiable*. Una vez que el sujeto cumple estas tres condiciones, ese sujeto es un sujeto moral. Es casi seguro que la posesión de habilidades meta-cognitivas expande el tipo y variedad de razones morales que un individuo es capaz de albergar. Sin embargo, cuando está cimentada en un mecanismo confiable, la sensibilidad normativa hacia las características que convierten una situación en buena – o mala – es suficiente para poseer razones morales básicas. Es en este sentido en el que Mishkin posee razones morales y, por tanto, cuenta como sujeto moral.

Es en este sentido, también, en el que Gracia puede poseer razones morales. Si las emociones de Gracia son apropiadamente sensibles a

las características que vuelven buenas – o malas – las situaciones, si esta sensibilidad es normativa, y si las emociones implicadas están producidas por un mecanismo confiable, esto es todo lo que se requiere para calificar a Gracia como alguien que actúa por razones morales. Si las pruebas empíricas que actualmente están siendo acumuladas por los etólogos cognitivos son explicadas mejor en los términos de estas tres características, entonces podríamos concluir legítimamente que Gracia, y muchos otros mamíferos sociales, son sujetos morales.

Notas

1. La frase que utiliza el autor es complicada de traducir: "We can attribute them by a way of a sentence embedded in a that-clause". Como la traducción "Podemos atribuírselo mediante la inclusión de una oración que contenga una oración subordinada-que" me parece bastante artificial, he optado por traducir la frase simplemente como "oración subordinada" (*N. del T.*).
2. El autor utiliza la expresión "moral subjecthood", para la cual creo que no hay una traducción directa en castellano. El término que he empleado, "subjetividad moral", debería entenderse como equivalente a la afirmación empírica de ser un "sujeto de experiencia moral", en modo alguno como un compromiso normativo del autor con el subjetivismo ético (lo cual puede o no ser el caso) (*N. del T.*).
3. Los términos que emplea el autor son "Moral Super Blindsight". Blindsight significa "visión ciega" (y, a veces, "ceguera cortical") y se emplea para referirse a la paradójica y poco frecuente condición neurofisiológica de una persona que no ve conscientemente porque su campo visual está dañado por una lesión (cortical) pero, sin embargo, conserva cierta sensibilidad visual residual y su conducta puede estar guiada por información visual (*N. del T.*).