

BERNARD WILLIAMS

# LA FORTUNA MORAL

Ensayos filosóficos 1973-1980

Traducción de

SUSANA MARÍN



INSTITUTO DE INVESTIGACIONES FILOSÓFICAS

*Colección:* FILOSOFÍA CONTEMPORÁNEA

*Director:* DR. LEÓN OLIVÉ

*Secretario:* DR. AMBROSIO VELASCO

UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO  
MÉXICO 1993

Título original:

*Moral Luck. Philosophical Papers 1973-1980*  
Cambridge University Press, 1981.  
Copyright © 1981, Cambridge University Press

All rights reserved. No part of this book may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or any information storage and retrieval system, without permission in writing form from the Publisher.

*A Isaiah y Aline Berlin*

Primera edición en español: 1993

DR © 1993, Universidad Nacional Autónoma de México

INSTITUTO DE INVESTIGACIONES FILOSÓFICAS

Circuito Mario de la Cueva  
Ciudad de la Investigación en Humanidades  
Coyoacán, 04510 México D.F.

Impreso y hecho en México  
ISBN: 968-36-3164-9

el modelo de las oraciones neutrales ha de ser incoherente.

La verificación de este tipo debe ser incoherente en relación no solamente con el modelo de las oraciones neutrales, sino con cualquier visión que intente ofrecer lo que puede llamarse una representación "absoluta" del mundo, en el sentido (bastante ambicioso, pero menos ambicioso que algunos otros sentidos que se le han dado a esta expresión) de una representación del mundo como es, a diferencia de la forma en que se presenta peculiarmente ante cualquier grupo en virtud de las peculiaridades de dicho grupo.<sup>14</sup> Hay quienes dudarán que sea posible lograr tal imagen absoluta del mundo, y en particular que pueda lograrse mediante la investigación científica. Pero quienes han deseado una filosofía centrada en un punto de vista científico del mundo no lo han dudado, sino que han basado su filosofía en una visión esperanzada de una imagen científica del mundo sólo porque pensaban, y con razón, que tal imagen era lo único que podría lograr una representación absoluta de las cosas. Se ha pensado, y ciertamente lo han hecho los positivistas, que el aditamento positivista de la verificabilidad se relacionaba con su objetivo de ser una filosofía de la visión científica del mundo. Pero si este argumento es correcto, la inclinación verificacionista del positivismo constituía, por el contrario, un obstáculo básico para llegar a ser este tipo de filosofía.

## ÍNDICE

PREFACIO . . . . .	7
RECONOCIMIENTOS . . . . .	11
1. PERSONAS, CARÁCTER Y MORALIDAD . . . . .	13
2. LA FORTUNA MORAL . . . . .	35
3. UTILITARISMO Y AUTOINDULGENCIA MORAL . . . . .	59
4. POLÍTICA Y CARÁCTER MORAL . . . . .	77
5. CONFLICTOS DE VALORES . . . . .	97
6. LA JUSTICIA COMO VIRTUD . . . . .	111
7. RAWLS Y LA APUESTA DE PASCAL . . . . .	123
8. RAZONES INTERNAS Y EXTERNAS . . . . .	131
9. <i>DEBER</i> Y OBLIGACION MORAL . . . . .	147
10. LA NECESIDAD PRÁCTICA . . . . .	159
11. LA VERDAD EN EL RELATIVISMO . . . . .	169
12. WITTGENSTEIN Y EL IDEALISMO . . . . .	183
13. OTRO TIEMPO, OTRO LUGAR, OTRA PERSONA . . . . .	207

<sup>14</sup> He intentado hablar más sobre esta concepción, su historia y su situación actual en *Descartes: The Project of Pure Enquiry*, Penguin, Harmondsworth, 1978.

más para qué sirve el concepto y, por la misma razón, cuántas otras características tuyas se pueden dar por sentadas.

El escepticismo sobre la moralidad como libre de la fortuna no puede dejar el concepto de moralidad donde estaba, y tampoco puede permanecer éste imperturbable ante el escepticismo sobre la muy cercanamente relacionada imagen que tenemos de la existencia de un orden moral, dentro del cual nuestras acciones tienen una relevancia que no puede adscribirles el mero reconocimiento social. Estas formas de escepticismo nos dejarán con *un* concepto de moralidad, pero ciertamente será menos importante de lo que normalmente consideramos al nuestro; y no será el nuestro, puesto que una característica particularmente importante del nuestro es la importancia que le adjudicamos.

### 3. UTILITARISMO Y AUTOINDULGENCIA MORAL

Mi problema surge de la cuestión de qué debe uno hacer en circunstancias en las que existen fuertes razones, particularmente de tipo utilitario, para hacer algo que se considera moralmente desagradable y en contra de lo cual tiene uno un fuerte compromiso personal. Por supuesto, también incluye la cuestión de lo que se dice y piensa sobre las acciones de otras personas en tales situaciones. Mi interés se refiere a una acusación que puede esgrimirse contra personas que rechazan acciones moralmente desagradables en tales casos, es decir, que son culpables de cierto tipo de autoindulgencia. Cuando la negativa del agente adopta la forma particular de decir que aunque sin duda otros harán mal, por lo menos esto no ocurrirá a través *suyo*, la acusación fácilmente puede adoptar la forma de decir que el agente muestra una actitud posesiva respecto a su propia virtud.<sup>1</sup>

El problema surge particularmente en relación con el utilitarismo. Si las razones para realizar la acción son suficientemente fuertes, desde un punto de vista utilitario, los utilitaristas dirían que el hecho de que la acción sea moralmente desagradable ciertamente no es una razón adecuada para no realizarla en este caso; como característica general de las acciones de este tipo, resulta en gran medida irrelevante respecto a cuestiones de qué hacer aquí y ahora, aunque puede resultar relevante para otros aspectos de la situación; así, podemos pensar bien del agente por sentir que este tipo de acción es desagradable, tomando su reacción como una señal que asegura su buen carácter. Dentro del contexto de una crítica al

<sup>1</sup> Esta frase aparece en una discusión al respecto de Jonathan Glover, *Proc. Arist. Soc.*, Volumen suplementario XLIX (1975).

utilitarismo he invocado anteriormente<sup>2</sup> la noción de *integridad* en relación con esto, y es dentro de este contexto como discutiré el problema aquí, es decir, tomando las razones que impelen a una persona a realizar la acción desagradable por razones utilitarias. Sin embargo, la estructura general de este problema para la acción individual no se limita a este tipo de casos, y espero que mi discusión ayude a revelar este hecho.

Existe un conjunto de problemas muy estrechamente relacionados con éste, que son los problemas políticos, tomando este concepto en un sentido amplio de acción en un dominio y en una esfera públicos, aunque los temas más claros e importantes surgen en asuntos de Estado. El más claro de todos los casos son las acciones realizadas por políticos en el ejercicio de su cargo en el contexto de asuntos internacionales, pero otros temas similares pueden llegar incluso a casos como las acciones de políticos que negocian para mejorar en su posición, sus aspiraciones y las expectativas morales de sus seguidores. No intentaré aquí discutir estos problemas de la moralidad política —el tema de las “manos sucias”.<sup>3</sup> Es importante que no solamente representan un caso especial del tema que discutiré, no sólo son ejemplos de este tema que surgen en el campo de la política. En los ejemplos políticos más claros tenemos dos características especiales. En primer lugar, el agente guarda una relación con otras personas —ciudadanos, seguidores, votantes, etc.— en la cual se supone que logra resultados que implican, y se puede saber que implican, tales acciones; y esta relación en sí puede tener una dimensión moral, por ejemplo, de confianza. Digo “se supone que logre resultados que implican...” en lugar de “se supone que realiza tales acciones” porque el público algunas veces adopta, y los medios de comunicación a menudo pretenden adoptar, un punto de vista moralizador mediante el cual se supone que los políticos no realizan las acciones requeridas para lo que se supone que deben lograr. En segundo lugar, la esfera de operación es en sí menos moralizada y menos estructurada por expectativas moralizadas que por lo menos muchas otras actividades, por lo menos en comunidades establecidas: las relaciones internacionales, por supuesto, representan el máximo ejemplo de ello.

<sup>2</sup> En J.J.C. Smart y B. Williams, *Utilitarianism: For and Against*, CUP, 1973.

<sup>3</sup> Véanse la discusión de Michael Walzer en *Philosophy and Public Affairs* (Invierno de 1973), y el cap. 4 de esta obra.

Estos dos factores son distintos entre sí. Los temas del segundo tipo pueden surgir incluso si no existiera nadie ante quien y hacia quien se fuese responsable: algunos moralistas tradicionales, aunque no todos, han pensado que existían restricciones en el grado hasta el cual las consideraciones morales se aplican en el estado de naturaleza, y creían en una analogía a *silent leges inter arma*. Nuevamente, la primera característica puede surgir sin la segunda, pero sin el segundo factor habría mayores dudas de que el papel ejercido por los responsables sea legítimo o aceptable; las expectativas que las personas tienen puestas en el líder se ven afectadas por su percepción del terreno sobre el cual los está guiando.

Por estas razones, las cuestiones de las manos sucias no solamente representan un caso especial de nuestro problema: o más bien, suponer que lo representan implica una exigencia crucial sobre su solución. Nuestro problema actual se refiere a la naturaleza y el contenido adecuado de lo que indudablemente constituye el juicio moral individual de una persona y aborda lo que ciertamente es una elección moral entre soluciones morales (dejando de lado un enfoque que realmente *definía* las consideraciones morales en términos del utilitarismo). La cuestión de las manos sucias, por lo menos en su forma más fuerte, concierne al papel que debe desempeñar en el asunto el juicio moral individual de una persona. El problema que nos concierne se refiere a la conciencia moral del individuo y la forma en que debe valorar la situación; la cuestión de las manos sucias plantea la cuestión de si su conciencia moral, y la forma en que valora la situación, no son simplemente irrelevantes.

Un tema que surge notablemente con respecto a ambas cuestiones, pero que, nuevamente, no discutiré, es el grado hasta el cual, y las formas en las que, las acciones ofensivas a la moralidad pueden justificarse retrospectivamente —incluso tal vez justificarse moralmente— por el éxito; y, si así es, qué puede considerarse como éxito. En su forma menos aceptable, éste es el punto de vista de que la historia puede justificar incluso las atrocidades políticas. No obstante, ni el bajo nivel de aceptabilidad de esta aplicación, ni (aún menos) cierta supuesta garantía ofrecida por el sentido del término “moral” nos debe conducir a subestimar este enfoque en general: hay más en él de lo que la gente quiere admitir. Pero éste es tema para otra ocasión.

Nuestro problema surge en casos en que el agente enfrenta una

elección segura entre una acción detestable y un resultado que utilitariamente sería peor: donde “una elección segura” significa que puede elegir entre realizar o no realizar la acción, y ciertamente queda fuera de toda duda razonable que si no realiza la acción, entonces se dará tal resultado, o algo incluso peor.<sup>4</sup> Existen argumentos conocidos que sugieren que no hay elecciones de este tipo, o muy pocas, que sean realmente seguras. Por una parte, los utilitaristas subrayan la importancia de los efectos secundarios para calcular un equilibrio en utilidad entre actuar o no hacerlo: cuando existen efectos secundarios, se dice que la acción desagradable tiene menos utilidad que a primera vista, y también puede tener menos utilidad que el otro resultado. Vale la pena notar que el nivel de probabilidad relacionado con estas condiciones generalmente permanece bastante indeterminado. Algunos de estos efectos, en los cuales los defensores del utilitarismo otorgan gran peso a lo abstracto, son tan problemáticos que en cualquier caso real un utilitarista consistente y razonable se vería obligado en gran medida a descartarlos. De cualquier manera, supondremos que estamos tratando sobre una clase de casos en los cuales, tomando en cuenta todas estas consideraciones, la balanza de la ventaja utilitaria favorece la acción que (de otra manera) resultaría detestable. Claramente, ningún utilitarista podría decir, y pocos desearían hacerlo, que no puede existir ningún caso de este tipo.

Otro plan de acción para hacer dudar de la confiabilidad de tales elecciones, esta vez utilizado por los antiutilitaristas, es sugerir que la eficacia de la acción detestable (por ejemplo, para prevenir grandes daños que ocurrirían de otra manera) es más dudosa que lo que supone el ejemplo. Ésta es una trayectoria que a menudo siguen quienes defienden una posición absolutista en casos de acciones detestables forzadas por amenazas de secuestradores y ese tipo de casos, en el sentido de que el carácter mismo de la amenaza demuestra que se pueden tener razones para dudar de la eficacia de ceder a ella: ¿por qué habría de esperarse que las personas que expresan las amenazas cumplan sus promesas? Como argumento

<sup>4</sup> Supuestamente se puede aceptar que mientras más horrible sea la acción que se ha de justificar por la expectativa de un buen resultado, más probable tendrá que ser que el resultado verdaderamente siga a la acción: supóngase que esto está incluido en el caso. Véanse dos ejemplos de este tipo en Smart y Williams, *op. cit.*, pp. 97-99.

*en general*, esto me parece, claramente, una escapatoria. Por supuesto existen algunos casos en que razonablemente se podría apostar que no se logrará nada cediendo a las amenazas, pero hay otros en los que tal apuesta no sería razonable, y simplemente es evasivo pretender que tenemos una seguridad *a priori*, aplicable a cualquier caso, de que resulta inadecuadamente indudable que la acción tendrá el efecto esperado.<sup>5</sup>

De cualquier manera, sólo hay ciertos tipos de ejemplos a los que se puede aplicar relevantemente este razonamiento, a saber, aquellos en los cuales, si las personas que producen las amenazas no cumplen, el resultado global es peor que si no se hubiera realizado la acción detestable. No todos los casos en que surge nuestro problema —ni siquiera todos aquellos donde existen amenazas— tienen esta estructura. Existe el caso en el que se me pide que mate a alguien y se me dice que, si me niego, otra persona lo matará, además de matar a otros. Si pensamos solamente en términos de resultados, entonces el único resultado concebible que realmente es mejor que los que implican que yo acepte, es aquel en el cual me niego y ellos deciden no matar a nadie; pero no existe ninguna probabilidad en absoluto al respecto. Si las otras personas hacen lo que en el caso del secuestrador correspondería a romper la promesa, es decir, si acepto y de todas maneras matan a los otros, entonces el resultado, considerado como resultado, solamente es igual al que seguramente ocurrirá si me niego. Así, en términos de resultados necesitamos solamente una probabilidad distinta de cero de que hagan lo que dicen para que mi aceptación sea racional.

En general, los argumentos de este tipo parecen demasiado inclinados a confundir la idea de que es posible algún factor ignorado en el ejemplo, con la idea de que tiene cierta probabilidad indeterminadamente alta. En esto se parecen notablemente a algunos argumentos que esgrimen sus oponentes utilitaristas. Sin duda la razón es la misma: cada uno está intentando encontrar a su manera un argumento relacionado con las consecuencias para un

<sup>5</sup> La idea que subyace a esto me parece ser la *unidad de los vicios*, un principio psicológicamente infundado. Una aplicación extraña de una idea muy semejante es un argumento adoptado de McTaggart por P.T. Geach, en el sentido de que no podríamos tener razones para creer en un infierno injusto: siendo la revelación la única base para creer en el infierno, no tenemos razón alguna para considerar confiable lo que nos comunique un Dios tan malo como para manejar un infierno injusto. Véase *Providence and Evil*, Cambridge, 1977, pp. 134-136.

sentimiento cuyas raíces no se encuentran para nada en las consideraciones respecto a las consecuencias.

Aceptemos la existencia de una elección segura del tipo en cuestión. Alguien que conscientemente sigue el curso antiutilitario en tal caso podría verse sujeto a ser acusado de preocuparse por su propia integridad o pureza o virtud a expensas de otros. Para utilizar una frase como título general, aunque no siempre resulte la mejor frase, llamemos a esto la acusación de autoindulgencia moral. Lo primero que quiero discutir son ciertas condiciones necesarias para que tal acusación resulte apropiada. Al hacerlo supondré que esta acusación no es, ni tiene intención de ser, sólo trivialmente *equivalente* a una afirmación desaprobatoria de que alguien, por razones morales, actuó conscientemente en forma antiutilitaria. Supongo que no se pretende una equivalencia, puesto que alguien que hace tal acusación de esta manera seguramente pretende ensalzar la solución utilitaria a tales casos, y por ende, indirectamente, el sistema utilitario, al colocar enfoques no utilitarios en algunas de sus aplicaciones bajo una acusación que tiene cierta fuerza independiente y que podría reconocerse ya como objeción. Esta fuerza independiente es la que trataré de revelar; y, más particularmente, supondré que la acusación de autoindulgencia moral imputa un tipo específico de *motivación*.

De hecho no es una condición suficiente ni necesaria para que esta acusación sea apropiada que el agente haga conscientemente y por una motivación de tipo moral algo que sea antiutilitario. No es suficiente considerando, por ejemplo, el caso de una persona que, haciendo valientemente lo que considera su deber (o incluso sólo valientemente), hace que lo maten durante un proyecto antiutilitario. Esta acción podría considerarse imprudente o tonta, pero no, con base en esto, *moralmente autoindulgente*: lo que contribuye a esto posiblemente podría incluir el hecho de que la persona misma paga un precio alto. Tampoco es una condición necesaria. Es posible que alguien esté sujeto a una acusación de autoindulgencia moral cuando las consideraciones morales que influyen en él son en sí utilitarias. Alguien podría incurrir en esta acusación en ciertos casos (no todos) si por razones de bienestar utilitario general, le fallara a alguien que dependía de él. Si la persona que se negó a disparar cuando se lo pidieron con amenazas estaba manteniendo sus manos limpias de lo que el utilitarista consideraría como suciedad finalmente irreal, *esta persona está manteniendo*

sus manos limpias de lo que, para el utilitarista, es suciedad real.

¿Qué impulsaría a alguien a aplicarle esta acusación a esta persona? Una característica podría ser que en realidad no parecían importarle mucho otros beneficiarios particulares. Esto no puede significar simplemente que no había beneficiarios particularmente identificados por los cuales se preocupaba; *ése* sería el caso, por ejemplo, de una persona que actuara honorablemente en interés de los habitantes no identificados de una ciudad identificada o, para tomar un caso más radical, que actuara para evitar peligros de radiación para las personas del futuro. Tampoco el refinamiento de esta idea nos conducirá al meollo de esta acusación; pero la idea sugiere algo que se acerca mucho más a él. Una cosa que puede expresar esta idea es la sospecha de que lo que le preocupa al agente no son tanto las demás personas, sino el hecho de que él mismo se preocupe por otros. Tiene una imagen de sí mismo como utilitarista virtuoso, y esta imagen es más importante en su motivación que cualquier preocupación por otras personas, en particular aquella persona hacia quien se le pide particularmente que muestre preocupación.

Este tipo de preocupación *reflexiva* es la que, yo sugiero, se encuentra relacionada significativamente con la acusación de autoindulgencia moral. Puede surgir con cualquier motivación moral. Así, una persona puede actuar por generosidad o lealtad, y actuar en forma antiutilitaria, y no provocar la acusación de autoindulgencia moral, pero tal acusación sí se daría si se sospecha que esta acción fue motivada por su preocupación por su propia generosidad o lealtad, o el mejoramiento o preservación de su propia autoimagen como persona generosa o leal. En el caso de una persona que actúa en forma antiutilitaria por razones de tipo no moral, la acusación de autoindulgencia moral en todo caso no se aplicaría, puesto que no es "moral" el tipo de autoindulgencia que busca, si es que acaso hay alguno. Pero existen contrastes fuertemente análogos en el tema de la reflexividad. Una cosa es que una persona actúe en forma antiutilitaria por su gran amor hacia Isolda, y otra que lo haga porque le preocupa su imagen de sí mismo como un gran Tristán. Esta distinción se aplica incluso al caso del egoísmo. Uno puede actuar egoístamente, es decir, de una manera motivada porque se desean cosas para sí mismo con indiferencia hacia el bienestar de otros, pero sería diferente actuar a partir de una concepción de uno mismo como persona que actúa de tal

manera. Aunque esto último probablemente no sea mejor, existen posibilidades de que sea un poco más grandioso.

Me imagino que en general existe una distinción conceptual clara entre la motivación de primer orden en cada uno de estos casos y su sustituto reflexivo de segundo orden. Sin embargo, después de este reconocimiento muy general existen muchos aspectos en los cuales incluso a nivel analítico, por no hablar de la realidad psicológica, las fronteras son bastante poco definidas. Haré unos cuantos comentarios respecto a lo que obviamente es un conjunto considerable de cuestiones.

Una condición necesaria para adscribirle a un agente la motivación de segundo orden es que también suponemos que posee el concepto de la motivación de primer orden en cuestión. Existe una distinción particularmente clara entre los dos tipos de motivaciones cuando es posible verse motivado en cierta forma moral sin poseer el concepto relevante de tal motivación en absoluto. Algunos tipos de motivación virtuosa permiten esto, y otra consecuencia errada de la teoría moral kantiana es que se considera que la única motivación moral genuina implica esencialmente que el agente esté consciente de este tipo de motivación. Pero incluso si el agente posee el concepto de cierta motivación virtuosa, pudiera ser que no lo aplique a su propio caso: en el espacio que otorga, con algunas virtudes, esta posibilidad, existe lugar para algo como una inocencia inteligente. E incluso si, por último, el concepto se aplica y existe el pensamiento de su propia disposición, no es lo mismo que si ese pensamiento proporciona su motivación. Un punto que vale la pena investigar más es que en el caso de algunas virtudes (como, tal vez, la valentía) la presencia de tal pensamiento puede resultar alentadora para la motivación de primer orden, mientras que en el de otras no es así, ya que la presencia de este pensamiento tiende a destruir la motivación de primer orden. Hasta el punto en que esto es así, habrá una razón (hay otras) por la cual algunas virtudes son accesibles sólo imperfectamente para los agentes reflexivos y concientes de sí mismos, al igual que existen otras virtudes que solamente son totalmente accesibles para ellos.

Muy bien podría ser que el camino para adquirir y conservar las motivaciones virtuosas de primer orden requiera un tipo de autoestima que puede implicar hasta cierto punto y de cierta forma las motivaciones de segundo orden. Es una cuestión para la teoría psicológica el grado hasta el cual esto ocurre, aunque este grado

ciertamente es limitado, por ejemplo por los asuntos de posesión de conceptos que se han mencionado. También es cuestión psicológica, tal vez menos de la teoría que de la observación común, hasta qué grado opera qué tipo de motivación. Nada de lo que estoy diciendo tiene la intención de implicar que es sencillo señalar hasta qué grado está operando qué tipo de motivación, ya sea en el caso de otra persona o —lo que en la naturaleza de este asunto es una cosa muy distinta— en el propio.

Empero, incluso si hay una diferencia entre estos tipos de motivaciones, aún existe una cuestión sobre qué es lo que se supone que tiene de malo la motivación de segundo orden, si es que algo tiene; particularmente, qué es lo que la convierte en autoindulgente. De hecho, algún filósofo podría argumentar que por lo menos en el caso de ciertos tipos de motivación de segundo orden, nada podrían tener de malo. Pues en las explicaciones que se han dado hasta ahora, parece que una persona podría verse motivada de alguna manera de segundo orden de este tipo si se preguntara a sí misma: “¿Qué haría yo si actuara como actuaría una persona generosa en este caso?”, y se viera motivada a actuar con base en la respuesta; y si produjera la respuesta correcta a la pregunta, y actuara con base en ella, entonces parece que haría exactamente lo que haría una persona generosa, y su razón no sería peor. ¿Constituye esto una autoindulgencia moral?

No, aunque como imagen de una deliberación moral este patrón ciertamente está muy distorsionado (ya sea que la distorsión se encuentre en la imagen o en lo que hay en ella). Lo que falta en esto para que sea un caso de autoindulgencia, por muy extraño que sea en otros aspectos, es algún elemento de autoestima, un punto sugerido por el hecho de que, después de todo, la generosidad de una persona hipotética idealmente generosa es lo que se invoca en este caso, y no la del propio agente. En este caso lo que nos puede engañar son frases como “se preocupa por ser generoso”. Esto puede significar meramente que se preocupa por hacer algo que sería generoso en el sentido en que eso es lo que le preocuparía hacer a cualquier persona generosa; o que se preocupa por conformar su conducta a cierto paradigma de una persona generosa, como el agente que acabamos de mencionar (este tipo de reflexividad parece, de hecho, como un ejemplo conocido de un desarrollo más primitivo, más que uno más completo); o puede significar que se preocupa por su propia generosidad, donde esto implica que



habría sustituido por un pensamiento de lo que se necesitaba, un pensamiento que enfoca en forma desproporcionada la expresión de su propia disposición, y que obtiene placer del pensamiento de que su disposición se habría expresado —más que obtener placer, como podría hacerlo el agente no autoindulgente, del pensamiento de cómo serían las cosas si actúa de cierta manera, siendo ésta (aunque no necesariamente tiene que pensarlo así) la expresión de su disposición.

Este tipo de reflexividad es el que produce el título de “autoindulgencia”. Implica una inversión en una división que me parece fundamental para cualquier vida moral o incluso cuerda, entre la preocupación por sí mismo y por los demás; implica una dirección errada no sólo de la atención, aunque esto también es cierto, sino genuinamente de la preocupación, y ambas dan como resultado diferencias en lo que se hace en realidad. Las distorsiones debidas principalmente a una atención desviada se conocen también en relación con las habilidades; en las que provienen de una preocupación desviada, las virtudes comparten con los afectos. Estas diferencias en lo que se hace encajan con algo que se mencionó anteriormente con respecto a la valentía, el peso de las evidencias que se le adscribe respecto a estas cuestiones al hecho de que el agente mismo pague un precio: *puede* hacer eso durante su observación reflexiva de su propia virtud, pero el espacio para ello se reduce.

Estos comentarios sobre la reflexividad y la motivación moral implican una cuestión difícil e importante que no se ha estudiado adecuadamente: cómo hemos de concebir la expresión de las disposiciones morales en el pensamiento deliberativo de un agente. Tenemos algunos puntos de vista en la filosofía sobre la referencia a las disposiciones para explicar y evaluar la conducta de otras personas. Tenemos algunos puntos de vista sobre la existencia de consideraciones morales en la deliberación práctica (aunque en gran medida se ven restringidas a cuestiones sobre la función en la deliberación de los “juicios morales”). Lo que parece hacernos falta es una representación coherente de algo que indudablemente es cierto, que las disposiciones morales claras, como la generosidad, se expresan en el contenido (y no solamente en las ocasiones) de las deliberaciones del agente. La única afirmación que hago sobre este tema aquí es que la expresión característica y básica de una disposición moral en la deliberación *no* es una premisa que se

refiera a tal disposición; no es la característica básica de las deliberaciones de una persona generosa que utilice la premisa “soy una persona generosa”. Sin importar lo que se diga después sobre este tema, esta negación ciertamente es correcta. Aunque una persona generosa se caracteriza en parte por lo que contienen sus deliberaciones, lo que contienen no son reflejos de su generosidad.<sup>6</sup>

Nos encontramos ahora en posición de observar mejor las relaciones entre el utilitarismo y la integridad en el tema de la autoindulgencia moral. Si la característica objetable de la autoindulgencia moral se identifica como cierto tipo de motivación reflexiva, entonces no puede contrastarse simplemente con la motivación utilitaria, pues, en primer lugar, puede contrastarse con muchas otras cosas, como, en general, con motivaciones virtuosas de primer orden. Además, la benevolencia utilitaria en sí está sujeta a esta deformación reflexiva. La razón por la cual la motivación utilitaria les parece a muchas personas el único enemigo de la autoindulgencia moral es que parece ser la expresión más pura de la preocupación por los otros en oposición a la preocupación por uno mismo; ¿no es el utilitarismo simplemente la expresión de la preocupación por todos, entre los cuales uno mismo es superado en número por los otros? Pero de hecho la distinción entre la preocupación por los otros y la preocupación por uno mismo de ninguna manera es igual a la distinción entre utilitario y no utilitario, y en el sentido en que las motivaciones de preocupación por otros que no son las del utilitarismo son susceptibles de deformación reflexiva para convertirse en un tipo de preocupación por uno mismo, también lo es la motivación utilitaria misma.

¿Qué ocurre con la preocupación por la propia *integridad*? Lo más sencillo que se podría decir al respecto sería que la integridad es un caso de virtud y que, al igual que otras virtudes, está sujeta a

<sup>6</sup> Y podemos agregar que no solamente son pensamientos tales como “necesita ayuda”; la aparición de tales pensamientos ciertamente distingue a unas personas de otras, pero poco hace por separar a las personas generosas de las no generosas. Tampoco lo hace el “juicio moral” “debería ayudar”; aparte de las bien conocidas cuestiones sobre la relación de esto con la motivación, no es especialmente una marca de la persona generosa tener que actuar con base en esa idea. Una respuesta probablemente tendría que empezar con la idea de que la representación básica en la deliberación de tal disposición adopta la forma “quiero ayudar...”; esto tiene la ventaja adicional de que no resulta ininteligible cómo se pueden comparar tales consideraciones morales en la deliberación con consideraciones bastante distintas.

deformación reflexiva. Pero creo que esto sería errado: más bien deberíamos decir quizá que la integridad no es una *virtud* en absoluto. Al expresar esto no quiero decir que no haya mucho que decir a su favor, como podría decirse que la humildad no es una virtud. Quiero decir que aunque es una propiedad humana admirable, no se relaciona con la motivación de la misma manera que las virtudes. No es una disposición que en sí produzca motivaciones, como la generosidad y la benevolencia; tampoco es una virtud de las a veces llamadas virtudes "ejecutivas", que en sí no producen una motivación característica, pero que son necesarias para la relación entre uno mismo y el mundo que le permite a uno actuar a partir de motivaciones deseables en formas deseables, el tipo que incluye la valentía y el autocontrol. Más bien, alguien que muestra integridad actúa a partir de las disposiciones y motivaciones que son más profundamente suyas, y también tiene las virtudes que le permiten hacerlo. La integridad no le permite hacerlo, ni es aquello a partir de lo cual actúa cuando lo hace.

Si esto es correcto, podemos ver por qué la integridad, considerada como virtud, puede tener un dejo de autoindulgencia moral. Pues si se considera como motivación, resulta difícil reconstruir su representación en el pensamiento excepto en la forma reflexiva objetable: el pensamiento tendría que ser acerca de uno mismo y el propio carácter, y del tipo sospechoso. Si la integridad tuviese que poseer un pensamiento característico, el pensamiento no podría basarse en nada excepto en uno mismo, pero no existe tal pensamiento característico, solamente los pensamientos relacionados con los proyectos en la realización de los cuales una persona puede mostrar su integridad. En forma relacionada, no podemos educar a alguien para que posea integridad, en el sentido de enseñarle a mostrarla o ejercerla; más bien uno provoca que se preocupe genuinamente por algo y tenga las características necesarias para vivir con base en ello.

¿Pero qué sucede con el pensamiento "no a través mío", el pensamiento de que incluso si otros van a provocar maldad e injusticia en el mundo, no será por agencia mía que esto ocurra?<sup>7</sup> Ciertamente esto es ya un pensamiento reflexivo, e implica por lo

<sup>7</sup> Blover, *op. cit.*, ha llamado a esto el "Principio de Solzhenitsyn" por una frase en el discurso que pronunció este escritor al recibir el premio Nobel. Hace bien en invocar el nombre, pero esta idea no debería considerarse como un *principio*.

menos un paso menos hacia la expresión simplemente no autoconsciente de las disposiciones antiutilitarias. Sin embargo, este pensamiento en sí no es un pensamiento motivante, y esas palabras no expresan ninguna motivación distinguible. No se trata simplemente de que no en todas ocasiones expresen una motivación particular. Más bien, en sí mismos no expresan ninguna motivación en absoluto: si uno está motivado a *no hacerlo uno mismo*, entonces existe alguna (otra) motivación que uno tiene para no hacerlo. Una motivación de este tipo es el miedo, y quizá en la forma particular del miedo a la contaminación puede atraer la acusación de cobardía hacia algunos agentes en el tipo de circunstancias que estamos discutiendo. Con la motivación del miedo en general, a menudo ocurre que el agente *preferiría ser capaz de hacer* lo que teme hacer. En relación con esto, el miedo a la contaminación es un caso especial que constituye o una excepción a ello o un caso particularmente complejo de ello. Pero de cualquier manera, el miedo, sea del tipo que sea, de ninguna manera constituye siempre la motivación de los agentes que emplean esas palabras.

Un caso bastante distinto, quizá limitante, de una motivación que subyace a esas palabras se relacionaría con el orgullo, la motivación de alguien de quien no es cierto que preferiría poder hacerlo; podría hacerlo, pero no quiere que se haga, y se niega a que lo obliguen a hacerlo dándole razones para hacerlo. Una motivación pura, sin apoyo, de este tipo difícilmente resultaría adecuada para los casos que nos conciernen; puesto que los intereses de partes inocentes se han introducido a las razones para actuar; esto sería, en gran medida, arrogancia arbitraria. Pero un pensamiento similar, aunque distinto, podría expresarlo el agente en nuestro caso: similar en cuanto a que registra una negativa a verse forzado por las amenazas, persuasiones o ejemplo de otros; es diferente en cuanto a que no solamente está afirmando su propia independencia y su propio derecho de negarse, sino expresando otras motivaciones que tiene para no realizar la acción.

Por supuesto, los utilitaristas debatirían su derecho de negarse, pero el asunto es que la afirmación del agente "no a través mío" no expresa, en tal caso, una motivación del tipo sospechoso de autoindulgencia. En sí misma no representa ninguna motivación en absoluto, y las motivaciones que pueden subyacerla incluyen algunas que, por diversas razones, son sospechosas y otras que no lo son. La reflexividad de esta expresión en sí misma no representa

ninguna motivación sospechosa, sino solamente la autoconciencia de la negativa, sin importar el motivo de la negativa.

Existen muchas y diversas formas de disposiciones, patrones de sentimientos y deseos, que pueden motivar a las personas a realizar actos antiutilitarios; algunas en sí son virtudes, algunas son proyectos, inclinaciones y compromisos más particulares. La cuestión a la que me dirijo finalmente es el lugar que puede darles el utilitarismo a tales disposiciones. Pueden admirarse o deplorarse de diversas maneras, cultivarse o desalentarse. Algunas verdaderamente pueden admirarse y fomentarse por lo que, remota y finalmente, son razones utilitarias en el sentido de que, indirectamente, la presencia de estas disposiciones en el mundo fomenta el bienestar humano. Me parece incorrecto tratar de reducir todas las cuestiones sobre la evaluación de tales disposiciones a consideraciones utilitarias, y esto resulta en realidad incoherente, puesto que no existe ningún punto de vista coherente del bienestar humano en sí que sea independiente de temas tales como lo que a la gente le preocupa, en un espíritu no utilitario, en relación con asuntos como estas disposiciones. Pero ese no es el tema que me preocupa por el momento, y si esta argumentación examina las disposiciones de este tipo, a las que puede otorgárseles un valor utilitario indirecto, entonces supuestamente tendrá cierta fuerza *ad hominem* contra el utilitarismo.

La dificultad yace en que tales disposiciones son patrones de motivación, sentimiento y acción, y no puede ser que al mismo tiempo el mundo contenga tales disposiciones, y que sus acciones cumplan con los requisitos del utilitarismo. Si alguien es una persona que genuinamente y de todo corazón posee algunas de estas disposiciones admirables, no puede ser también alguien en cuyos pensamientos y acciones se reflejen invariablemente los requisitos del utilitarismo, y a nadie le gustaría ser así. Si alguien desea que el mundo esté lleno de personas generosas, cariñosas, vigorosas, resueltas, creativas y verdaderamente felices, no desea que haya personas que piensen uniformemente de manera que sus acciones satisfagan los requisitos del utilitarismo.

La suposición de que se podrían combinar las disposiciones que se desean y admiran en el mundo con acciones que satisfacen al máximo los criterios utilitarios surge de varios errores. Uno de ellos es una idea que el utilitarismo, aunque la niega, está de hecho dispuesto a compartir con otras imágenes de la experiencia moral,

y de hecho de la racionalidad práctica en general: que los procesos del pensamiento práctico son trascendentales para la experiencia y no ocupan en realidad ningún espacio psicológico. Pero de hecho, pensar de una manera y no de otra con respecto a lo que se va a hacer implica ser empíricamente diferente, ser cierto tipo de persona, y no es posible combinar todo tipo de reflexión con todo tipo de disposición. Los utilitaristas ignoran esto hasta cierto punto a nivel del individuo, pero se han especializado en ignorarlo a nivel social, suponiendo, por ejemplo, que pudiera haber una élite de pensadores utilitaristas que poseyeran una doctrina esotérica desconocida para otros, sin que se hubiese especificado ninguna forma de organización que convirtiera esta estructura en una realidad social.

En segundo lugar, existe el error, también compartido con otros, de disociar los pensamientos y las decisiones morales de los sentimientos morales. Es una creencia común que hay una forma de debilidad que consiste en verse sobrecogido por sentimientos morales no estructurados y que hay otra que es un tipo de melindrosidad. A menudo éstas se deben a confusión, a falta de conocimiento de uno mismo. Pero la cura no puede o no debe encontrarse en enseñarles a las personas a hacer a un lado sus sentimientos morales, a disociarse de ellos.<sup>8</sup> A los teóricos que apoyan esta idea les atraen casos tales como el del católico caído que ha perdido la fe pero se siente culpable si no va a misa. Pero, sin importar lo que haya que decir sobre ese caso, no puede ser un paradigma de lo que el utilitarista requiere. El católico caído intenta disociarse enteramente de la misa y sus exigencias, alcanzar una posición en la que ningún sentimiento de ese tipo tenga significación alguna. Pero nada de esto se aplica a la persona con sentimientos antiutilitarios en un caso como el que estamos discutiendo. Estos sentimientos representan algo que él en general apoya, y que el utilitarista,

<sup>8</sup> De hecho, necesitamos aquí una teoría de los sentimientos morales. Se podría sugerir un enfoque a las cuestiones de disociación de los sentimientos morales por cierto contraste entre el sentimiento moral y la percepción de los sentidos. Las perspectivas, de tipo racionalista, que abogan con más fuerza por la disociación de las sensaciones perceptuales, por lo menos subrayan una verdad: que la finalidad del conocimiento objetivo es disociar los pensamientos sobre el mundo, ciertamente de lo que claramente es uno mismo y quizá (según las perspectivas realistas) de cualquier cosa que sea distintivamente humana. Pero ése no puede ser el objetivo de los pensamientos y la experiencia moral, que deben implicar primero la comprensión del mundo de tal manera que se pueda, como ser humano particular, vivir en él.

suponemos, en general quiere que apoye.

Nadie está sugiriendo que los sentimientos morales deban expresarse sin ser modificados por el pensamiento (en un extremo, esta idea ni siquiera resulta comprensible). Existen, además, ciertos sentimientos morales, en particular los relacionados con el cumplimiento de las reglas, que pueden formarse por la experiencia de maneras que hasta cierto punto caben y tienen lugar para los pensamientos utilitarios: hasta cierto punto, esto ocurre con las reglas de cumplir con las promesas y decir la verdad. Pero no existe razón alguna para creer, en el caso de muchas disposiciones del tipo que resulta deseable tener en el mundo, que pueden conservar su posición e importancia y, no obstante, dar paso, sistemáticamente y siempre que se requiera, a las deliberaciones del pensamiento utilitario, haciendo objetos de disociación a los sentimientos relacionados con la disposición.

En relación con esto, no hay mucho que pueda obtenerse de una tercera línea de pensamiento, que también puede fomentar una visión demasiado simple de estos problemas: la distinción supuestamente clara entre juzgar la acción y juzgar al agente. Si una persona tiene una disposición del tipo que es bueno que tenga, y si lo que hizo fue solamente lo que una persona con tal disposición haría en tal caso, pero (como afirmo que debe ocurrir algunas veces) era antiutilitario: ¿con qué fuerza se puede decir que lo que hizo fue de hecho incorrecto? Es importante el hecho que *no* haya la fuerza —que en realidad le daría cierto sentido a la distinción entre la acción y el agente— de que, si se hubiera encontrado en una posición en que pudiera haber realizado mejor sus deliberaciones, habría actuado de otra manera. Realizó sus deliberaciones tal como lo hace una persona así, y es bueno que sea una persona así. Por la misma razón, no puede significar que deberíamos tratar de educar a las personas de tal manera que no cometan tales errores. Si tiene algún contenido decir que esta persona hizo algo incorrecto, debe ser compatible con nuestra idea de que es bueno que las personas no siempre hagan lo correcto; y no solamente en el sentido muy general en que podríamos reflexionar sobre los inesperados beneficios que pueden surgir incluso de acciones terribles, sino en el sentido más íntimo de que queremos que en el mundo haya personas que cuando se pregunten “¿Qué sería correcto hacer?” den, en tipos de ocasiones definitivamente especificables, la respuesta incorrecta.

La teoría del utilitarista, una vez que admite el valor de estas disposiciones, lleva la pregunta “¿Qué sería correcto hacer?” muy lejos de la pregunta “¿Qué respuestas es deseable que la gente esté dispuesta a dar a la pregunta ‘¿Qué sería correcto hacer?’” La tensión creada por esta separación es muy grande, y existe una presión muy fuerte, si el utilitarismo ha de conservar una identidad identificable *dentro* del pensamiento moral, para que rechace o diluya irremediadamente el valor de estas otras disposiciones, en una regresión a la imagen del hombre que ofrecía francamente el utilitarismo original, en el cual, idealmente, el hombre tiene solamente proyectos privados o sacrificables de alguna otra forma, junto con la única disposición moral de la benevolencia utilitaria. Espero haber mostrado que no podemos aceptar esa falsa imagen rechazando otras motivaciones morales, en sus apariciones antiutilitarias, como ejemplos de autoindulgencia moral.