

Colección Teorema

Bernard Williams

*Introducción a la ética*



EDICIONES CÁTEDRA, S. A. Madrid

Título original de la obra *Morality. An Introduction to Ethics*  
Traducción de Manuel Jiménez Redondo

CLASIF. BJ15  
W5518  
FE 26.7.26  
FECHA 30-Junio-93  
PROCED. Lib. Paseo  
FACT. No. 355  
52/93



Cubierta de José Flores



© 1972 by Bernard Williams  
Ediciones Cátedra, S. A., 1982  
Don Ramón de la Cruz, 67, Madrid-1  
Depósito legal: M.15912-1982  
ISBN: 84-376-0337-4  
Printed in Spain  
Impreso en LAVEL. Los Llanos, nave 6 - Humanes  
Papel: Torras Hostench, S. A.

# BIBLIOTECA UNIVERSIDAD NACIONAL

## Índice

Prefacio . . . . .	11
El hombre amoral . . . . .	17
Subjetivismo: consideraciones iniciales . . . . .	27
Interludio: relativismo . . . . .	33
Subjetivismo: otras consideraciones . . . . .	39
«Bueno» . . . . .	51
Bondad y roles . . . . .	61
Estándares morales y la nota distintiva del hombre . . . . .	69
Dios, moralidad y prudencia . . . . .	77
¿Sobre qué versa la moralidad? . . . . .	87
Utilitarismo . . . . .	95

26728



Instituto de Investigaciones Filosóficas  
BIBLIOTECA  
"DR. EDUARDO GARCÍA MAYNEZ"  
CIUDAD UNIVERSITARIA  
MEXICO 10 D. F.

es una creación humana.) Pues los hombres, como cuestión de hecho, encuentran valores en cosas tales como la sumisión, la entrega, la incertidumbre, el riesgo, incluso en la desesperación y en el sufrimiento, y estos valores difícilmente pueden ponerse todos en relación con el ideal central de la *felicidad*. Y si encontramos algunas explicaciones, psicoanalíticas quizás, o incluso zoológicas, de esas actitudes, y llegamos a considerarlas como aberraciones que hay que tratar de evitar, entonces ciertamente estamos cambiando el mundo desde el punto de vista de una cierta moralidad, y no meramente haciendo el mundo más sensible a lo que la moral incuestionablemente es.

«Bienestar» era nuestro punto de partida; después nos hemos ocupado de su comprensión como «felicidad». Tal vez cabría decir que pese a que algunas clases de concepciones morales rechacen la felicidad como noción central, existe todavía una noción de bienestar, mucho más amplia y no por ello sin contenido, en relación con la cual no se produciría ese rechazo. Es una cuestión que tiene sentido, pero yo no sé cómo responderla. Por un lado, los casos más extremos parecen dejarnos con una noción de bienestar que realmente no se encuentra a mucha distancia de «ser como los hombres deben ser», y ahí no queda contenido alguno. Por otro lado, al caracterizar esas posiciones, uno habla de lo que los hombres, de hecho, encuentran valioso, o necesitan, o desean; y si alguien dijera —bastante oscuramente, por cierto— que los hombres necesitan un mundo en el que haya riesgo, incertidumbre y la posibilidad de desesperarse, entonces, de una moralidad que subrayara esto en contraste con las moralidades que quieren dejar todo lo más sosegado posible, habría que decir todavía que se refiere al bienestar de los hombres. Sin embargo, algo queda todavía excluido por el uso de este término: sistemas de valores o de preceptos que no prestan ninguna atención en absoluto a lo que podemos entender que los hombres necesitan o desean.

## Utilitarismo

Al discutir la cuestión de si todas las concepciones morales se refieren en última instancia a la felicidad humana, no he supuesto que la cuestión sea la misma que la de si todas las concepciones morales tienen que ser en una u otra medida una versión del *utilitarismo*. Obviamente no se trata de la misma cuestión si tomamos el utilitarismo en el sentido estricto de que sólo hay un principio moral, el de buscar la mayor felicidad para el mayor número; que «felicidad» significa aquí placer y ausencia de dolor; y que este único principio moral (pues se trata de un principio moral) hay que aplicarlo a cada situación individual («utilitarismo del acto»). Obviamente, existen muchas formas en que una moralidad puede estar en última instancia referida a la felicidad humana sin por ello tener que identificarse con *eso*. Pero creo también que existen multitud de formas en que la moralidad puede en última instancia referirse a la felicidad humana sin identificarse con el utilitarismo, ni incluso tomado en un sentido más lato.

Una dificultad que surge al discutir este asunto es la falta de acuerdo en la cuestión de con qué extensión puede usarse el término «utilitarismo» sin que deje de ser adecuado o significativo. El término ha sido utilizado a veces para referirse a concepciones morales que nada en absoluto tienen que ver especialmente con la felicidad o con el placer; en este sentido se usa para referirse a cualquier concepción que sostiene que la rectitud o no rectitud de una acción depende siempre de las con-



secuencias de la acción, de su tendencia a conducir a estados de cosas intrínsecamente buenos o malos. Este sentido tan amplio —que probablemente quedaría reflejado mejor por el término «consecuencialismo» que por el término «utilitarismo»— no es el que vamos a discutir aquí; aquí nos interesan solamente las concepciones de esta clase que consideran expresamente la *felicidad* como la única cosa intrínsecamente buena, la que, se supone, tienen como objetivo las acciones y las instituciones sociales. Pero esa restricción todavía deja bastante espacio para diferentes clases de utilitarismo.

Discutir en el vacío qué es lo que podría o no podría contar como una forma reconocible de utilitarismo sería un ejercicio puramente verbal y sin interés. La cuestión solamente puede abordarse preguntando qué es lo *esencial* en la concepción utilitarista de la moralidad; y esto es algo que no puede descubrirse meramente, ni tampoco principalmente, consultando lo que tenían en la mente Bentham y J. S. Mill y otros exponentes clásicos del sistema, sino considerando cuáles son los atractivos de la concepción utilitarista para el pensamiento moral. A mi juicio son cuatro principalmente, lo que no significa negar que no se relacionen entre sí de una forma digna de examinarse. En primer lugar, es no-trascendental, y no hace apelación a nada fuera de la vida humana, y en particular no hace apelación a consideraciones religiosas. En particular, contribuye así a apoyar la demanda enteramente razonable de que la moralidad habría de quedar hoy, como cosa obvia, liberada del cristianismo. Puede incluso parecer que contribuye a apoyar —a causa de cierto conservadurismo que consideraré más tarde— una pretensión mucho menos razonable, y que Nietzsche se dio cuenta muy bien de que era una imbecilidad, que la moralidad así liberada del cristianismo habría de seguir siendo la misma que la antes vinculada al cristianismo. En manos más radicales, sin embargo, el utilitarismo promete cambios más radicales.

Segundo: su bien básico, la felicidad, parece mínimamente problemático: por muchas que sean las cosas en que las personas disienten, sin duda que todos por

lo menos desean ser felices, y tener como objetivo la mayor felicidad posible tiene sin duda que ser, cualesquiera sean las otras cosas que hayan de someterse a ésta, un fin razonable. Ahora bien, existe un notable problema en este punto sobre la transición del objetivo, supuestamente indiscutible, de buscar la propia felicidad, al objetivo; más discutible, de buscar la felicidad de otra persona, y el desafortunado Mill ha sido repetidas veces llamado al orden por los críticos por tratar (se dice) de llevar a cabo esa transición por medio de un argumento deductivo. Por mi parte dudo de que fuera esto lo que Mill trataba de hacer, pero en cualquiera de los casos el problema no tiene una fuerza especial contra el utilitarismo —no hay ninguna razón para suponer que el utilitarismo, más que ninguna otra teoría, tuviera que tener una fórmula mágica para sacar de su amoralismo al hombre amoral. Lo *esencial* es más bien que el utilitarismo, lo mismo en este que en otros aspectos, es una moralidad de *compromiso mínimo*: supuesto sólo el requisito mínimo para estar en el mundo moral, a saber, la voluntad de tomar en consideración los deseos de los otros tanto como los propios, el utilitarismo puede moverse ya en ese terreno. Una cuestión mucho más interesante es la de si el «indiscutible» objetivo de la felicidad puede de hecho ser puesto al servicio de los propósitos utilitaristas. En la sección anterior ya hemos visto alguna razón para dudar de que la felicidad tenga que ser el objetivo de la vida humana; pero incluso dejando de lado esas cuestiones, no está en absoluto claro que cualquier sentido en el que (más o menos) quepa admitir un fin tal como indiscutible, sea también un sentido en favor del cual pueda hacerse trabajar al utilitarismo. Esta es una cuestión central: estaremos en mejor situación para abordarla cuando hayamos considerado los atractivos tercero y cuarto del utilitarismo.

El tercer atractivo es que los asuntos morales pueden, en principio, decidirse por un cálculo empírico de las consecuencias. El pensamiento moral se hace empírico y, en cuestiones de gestión pública, se convierte en un asunto de ciencia social. Y esto siempre ha sido conside-

rado por muchos como uno de los rasgos más gratificantes del utilitarismo. No es que se piense que los cálculos son fáciles, o incluso prácticamente posibles en algunos casos; el encanto radica más bien en que la naturaleza de la dificultad pierde, por lo menos, todo su misterio.

En cuarto lugar, el utilitarismo proporciona una moneda corriente universal de pensamiento moral: los distintos intereses de las distintas partes, y las distintas clases de demandas que actúan sobre una de las partes, pueden todos (en principio) ser convertidos a la misma moneda, en términos de felicidad. Este supuesto, y esto es importante, tiene como consecuencia que una cierta clase de conflicto, bien conocido para algunas otras concepciones morales, es aquí imposible, a saber, el conflicto entre dos pretensiones que son válidas las dos, y sin embargo, irreconciliables. Bajo algún otro sistema, un hombre puede verse en una situación en la que (así le parece a él) cualquier cosa que haga implica que está haciendo algo malo. Para el utilitarismo esto es imposible. Las distintas exigencias a las que ese hombre puede sentirse sometido pueden ser reducidas todas a la medida común del principio de la mayor felicidad, y no puede haber otra idea coherente de la acción correcta o de la acción no correcta que la de *lo que es mejor hacer tomando todo en consideración*: y si dos formas de acción resultan iguales en este sentido, entonces realmente no importa por cuál de las dos decida uno. En contraposición con esto, muchas personas pueden pensar que una determinada forma de actuación era la mejor que podía adoptarse tomando todo en consideración en las circunstancias concretas del caso, pero que, no obstante, el actuar de esa forma implicaba hacer algo no correcto. Esta es una idea que, a mi juicio, en última instancia, al utilitarismo tiene que resultar incoherente. Esta es una de las razones por las que puede decirse (y ciertamente con razón) que para el utilitarismo la tragedia es imposible; pero el utilitarismo tiene también consecuencias más amplias, si no más profundas, que esa.

El utilitarista podría ciertamente volverse sobre este tipo de ideas y, después de reflexionar un poco, invocar

cosas tales como las consecuencias socialmente deseables de que la gente sea un poco aprensiva en relación con ciertas acciones, aun cuando esas acciones, en las circunstancias del caso, sean las mejores que pueden emprenderse: más tarde volveremos sobre este tipo de argumentación. Pero lo que como utilitarista está obligado a hacer es considerar como objetivo indiscutible del pensamiento moral la reducción del conflicto, la eliminación, siempre que sea posible, de los conflictos de valores que no sean solubles sin secuelas. Aquí, como en cualquier otro sitio, su preocupación es la eficiencia: la generación del conflicto es un signo de ineficacia de un sistema de valores y el utilitarismo tiene un artificio general para eliminarlos o resolverlos. Pero alguien podría preguntarse si tal eficiencia es un objetivo indiscutible. Uno puede ciertamente reducir el conflicto, o hacer la vida más simple, reduciendo el rango de las demandas que está dispuesto a tomar en consideración; pero en ciertos casos esto podría ser considerado no como un triunfo de la racionalidad sino como una evasión cobarde, una negativa a ver lo que hay que ver (y aquí cabría preguntarse una vez más si el subjetivismo desactivado deja realmente todo donde estaba).

Por atrayente que pueda resultar en general el utilitarismo, este cuarto atractivo suyo puede ser puesto en tela de juicio por razones importantes. Otras dificultades se agolpan cuando uno considera lo que presupone. Pues sólo podremos ser capaces de utilizar el principio de la mayor felicidad como medida común de las pretensiones o demandas de todos y de cada uno, si la «felicidad» de que se habla es en algún sentido *comparable* y en algún sentido *aditiva*. Solamente si podemos comparar las felicidades aquí implicadas para las diferentes personas y sobre los diferentes resultados, y también juntarlos en alguna suerte de felicidad general, podremos hacer que la cosa funcione. En un plano técnico estos problemas han sido objeto de disciplinas tales como la economía del bienestar y la teoría de la preferencia, que han trabajado en el marco de supuestos muy artificiales y que sólo han tenido un moderado éxito a la hora de ela-



borar esos problemas para la teoría económica. Aquí nos ocupamos de dificultades más generales. Si la «felicidad» implicada ha de ser tal como para permitir al utilitarismo cumplir sus promesas tercera y cuarta, ¿puede ser también el indiscutible objetivo que se prometía en la segunda?

La respuesta a esto parece que es que no. Bentham ofreció una explicación de la felicidad en términos de placer y ausencia de dolor, de la que muy claramente se suponía que cumplía todas las promesas a la vez; pero aun cuando hubiera satisfecho (cosa que no hacía) las condiciones de ser calculable, comparable y aditiva, no cumplía la condición de ser un objetivo indiscutible: cuanto más parecía la clase de placer que concebiblemente podía ser tratado en esos términos cuasi matemáticos, tanto más perdía el aspecto de algo a lo que fuera evidente que cualquier hombre racional tenía que tender, cosa de la que, aunque a disgusto, se dio cuenta Mill. Si, por un lado, la concepción de la felicidad se hace lo suficientemente generosa como para incluir todo aquello a lo que pudiera razonablemente tenderse como a una vida satisfactoria o a un ingrediente de esa vida, entonces progresivamente va perdiendo el aspecto de algo que pudiera ajustarse a las condiciones tercera y cuarta. Aparte de otras cosas, existe la dificultad de que mucho de lo que la gente incluye de hecho en el contenido de una vida feliz son cosas que esencialmente implican otros valores, como son, por ejemplo, la integridad, o la espontaneidad, o la libertad, o el amor, o la auto-expresión artística; valores que no solamente no pueden ser tratados de la forma que las condiciones tercera y cuarta parecen imponer a la «felicidad» del utilitarismo, sino que además parece, por lo menos en el caso de algunos de ellos, que se da de hecho una contradicción en la tentativa de pensarlos como algo que podría ser tratado de esa manera.

Esta es, pues, la primera dificultad general con que se enfrenta el utilitarismo. Su «felicidad» tiene que satisfacer ciertas condiciones si es que se quiere mantener lo esencial del utilitarismo; y la condición de que indis-

cutiblemente habría de ser el objetivo de la aspiración humana entra en conflicto con las otras codiciones que tiene que satisfacer para poder ser tratada como el utilitarismo exige que se la trate. Enfrentado con esta dificultad general, una de las formas en que el utilitarismo tiende a reaccionar es la de poner en tela de juicio los valores implicados en las concepciones más intratables de la felicidad, como irracionales, tal vez, o como reminiscencias de una época pasada. Tales argumentos pueden ofrecer sin duda ciertos puntos interesantes, pero su estrategia global es vergonzosamente circular: la racionalidad utilitarista queda convertida en test de lo que ha de contar como felicidad, para quitar del paso toda otra clase de felicidad que puede constituir una objeción contra el utilitarismo. Todo lo que se necesita para enfrentarse a esto en el plano teórico es una firme negativa a sentirse intimidado.

El problema, sin embargo, no se reduce al nivel teórico: se presenta también, y drásticamente, en la sociedad, y la negativa a sentirse intimidado puede ser aquí inadecuada o difícil de hacerse valer. En los casos de planificación, conservación, bienestar, y decisiones sociales de todas clases, un conjunto de valores que, por lo menos en teoría, son cuantificados en términos de recursos, se enfrentan con valores que no son cuantificables en términos de recursos: tales como el valor de conservar una parte antigua de la ciudad, o de atender tanto a la dignidad como al confort de los pacientes en una unidad geriátrica. Una y otra vez, los defensores de tales valores se enfrentan con el dilema, o bien de negarse a cuantificar el valor en cuestión, en cuyo caso desaparece también por completo de la suma, o bien de tratar de asignarle alguna cantidad, en cuyo caso distorsionan el asunto de que se trata y por lo general pierden también la discusión, ya que el valor cuantificado no es suficiente para inclinar la balanza. En tales asuntos, no es que los utilitaristas estén comprometidos a pensar que esos otros valores no importan; o que estén confinados a considerar valiosas las cosas que de hecho son susceptibles de ser tratadas por medio del análisis coste-beneficio. Ni

tan siquiera, quizás, están obligados a pensar que todo valor social deba, en un caso dado, poder ser tratado en términos parecidos a los del análisis coste-beneficio: podrían decir que no tienen por qué suscribir la idea de que la medida común de la felicidad sea el dinero. Pero están obligados a sostener algo que en la práctica tiene esas implicaciones: que en última instancia no existen valores inconmensurables. Pues no es precisamente un rasgo accidental de la concepción utilitarista el que la presunción sea siempre en favor de lo cuantificable, y que los otros valores se vean por fuerza en el dilema apoloético al que acabamos de referirnos. No es un accidente, porque (por una parte) el utilitarismo es, cosa nada sorprendente, el sistema de valores de una sociedad en la que los valores económicos son supremos; y también, en el plano teórico, porque la cuantificación en dinero es la forma obvia de aquello sobre lo que el utilitarismo insiste, la conmensurabilidad de los valores.

Existe una gran presión para la investigación de técnicas que hagan conmensurables esferas cada vez más amplias de valores sociales. Parte del esfuerzo debería dedicarse más bien a aprender —o tal vez, a reaprender— a pensar inteligentemente sobre conflictos de valores que son inconmensurables.

Me he referido a dificultades relativas al cumplimiento de las condiciones que el utilitarismo pone a la «felicidad». El hecho de que esas dificultades existan no significa en modo alguno que en un caso dado o en una determinada clase de casos no podamos tener una idea de cuál sería la solución utilitarista o de qué clases de cosas tendría en cuenta el utilitarista. Algo iría mal si esa conclusión fuera correcta, pues está claro que a veces podemos hacer esas cosas: al discutir el utilitarismo estamos discutiendo *algo*, y algo que con mucha frecuencia es perfectamente reconocible. Así pues, admitamos que, por lo menos en algunos casos, sabemos qué es lo que quiere decirse con analizar qué actuación conduciría a la mayor felicidad. Admitido esto, nos encontramos entonces con dos nuevas dificultades. Una es que el proceso de analizar tales consecuencias es. él

mismo, una actividad, una actividad que en las diversas circunstancias posee grados diversos de utilidad; y esto tiene que incluirse en la suma. La otra es que la respuesta a que se llega por el cálculo utilitarista del caso particular parece en ciertas ocasiones que es una respuesta moralmente incorrecta. Existe una clase de utilitarismo, denominado el utilitarismo de reglas (*rule-utilitarianism*), cuyo objetivo es precisamente resolver estas dos dificultades al mismo tiempo, haciendo uso para ello de un mismo artificio.

El primer problema es que cualquier cálculo utilitarista efectivo habrá de tener lugar en condiciones de considerable incertidumbre y de información muy parcial, y así es probable que sus resultados no sean fiables. Además, el proceso de cálculo mismo lleva su tiempo; y el estar dispuesto a hacer cálculos en cada caso tiene rasgos psicológicos que, como cuestión de hecho, pueden ser un obstáculo para cosas que pueden ser deseables desde un punto de vista utilitarista, como puede ser una actuación rápida y decidida. Siendo así las cosas, cabe pensar que en la práctica podrían seguirse mejores consecuencias si los agentes, más que hacer cálculos para cada caso particular, suscribieran ciertas reglas que, sin cálculo alguno, se aplicaran usualmente a los casos particulares; sería la adopción de esas reglas y no la elección de acciones particulares lo que se examinaría recurriendo al principio de la mayor felicidad.

Esta misma idea se invoca para explicar el otro hecho, que, de otro modo, constituiría una dificultad, a saber: que fácilmente podemos imaginar casos —por ejemplo, el caso en el que la condena de un inocente fuera necesaria y suficiente para evitar grandes males— en los que el resultado utilitarista parece que entra en conflicto con lo que muchos considerarían que es la respuesta moralmente correcta; y al igual que la justicia en este caso, también el cumplimiento de las promesas y el decir la verdad presentan dificultades cuando se interpretan desde el punto de vista del utilitarismo del acto. Se espera que el utilitarismo de la regla pueda disolver estas dificultades pretendiendo que todo lo que tiene que demos-



trarse es que estas reglas o prácticas de justicia, de cumplimiento de las promesas o de veracidad poseen una utilidad positiva sobre sus alternativas.

Esto es sólo un bosquejo. Muchas cosas que difieren entre sí en aspectos importantes pueden caer bajo la denominación de «utilitarismo de reglas», y también son diferentes las cosas que hay que decir sobre ellas<sup>5</sup>. Todo lo que voy a tratar de hacer aquí va a ser sugerir un par de puntos en relación con la cuestión de hasta qué punto el utilitarista puede ir consistentemente en la dirección de las reglas; y sostener que, o bien no puede ir lo suficientemente lejos como para resolver la segunda dificultad, o si no, que tiene que ir tan lejos que (tanto el utilitarista como cualquier otro) dejará de ser utilitarista.

Es ciertamente posible para un utilitarista, sin incurrir por ello en incoherencia, adoptar una práctica general para tratar una determinada clase de casos, aun cuando algunas aplicaciones particulares de la práctica conduzcan a un resultado diferente del que se habría obtenido por medio de un cálculo individual de esos ejemplos. El paradigma de esto es el sistema de contabilidad que se aplica para muchos servicios públicos, sistema por el que, ocasionalmente, puede enviarse una factura de un dólar aun cuando procesar cada factura tiene un coste de dos dólares: la razón es que resulta más barato enviar todas las facturas en el momento debido, por pequeño que sea su importe, que interrumpir el procesamiento de todas para extraer unas cuantas. Vamos a llamar a esto el «modelo del recibo del gas».

Ahora bien, el modelo del recibo del gas se refiere a consecuencias de hecho: las consecuencias que tiene de hecho el limitarse a aplicar una regla, por un lado, y las consecuencias que de hecho tiene una interferencia particular, por otro. Esta clase de modelo, por lo menos, no puede hacer más digerible para un utilitarista coherente una forma de argumentación que no invoca ni

<sup>5</sup> Para una detallada y sutil discusión de estos asuntos, véase el importante libro de David Lyon. *The Forms and Limits of Utilitarianism*, Oxford, Clarendon Press, 1969.

las consecuencias que de hecho tiene una elección particular ni las consecuencias que de hecho tiene el que todos se atengan a una regla, sino las consecuencias *hipotéticas* del ajuste *imaginario* a una regla. Así, el modelo familiar de argumento moral consistente en la pregunta «¿Qué ocurriría si todos hicieran eso?» no puede tener ningún efecto sobre un utilitarista consistente, a menos que su acción hubiera de tener el efecto de que todo el mundo hiciera eso, cosa que por lo general es bastante implausible. Una consecuencia puramente imaginaria no podría figurar más en un cálculo utilitarista de lo que puede hacerlo la felicidad o infelicidad de personas puramente imaginarias. Así pues, el modelo del recibo del gas, por lo menos, no puede ponernos en el camino de esa clase de argumentación por generalización.

Si el utilitarista desea justificar el uso de una argumentación por generalización que se refiera a consecuencias imaginarias, tendrá que alejarse un paso más de las consecuencias que de hecho tienen las decisiones particulares, y tratar el asunto en términos de las consecuencias que *de hecho* tiene que la gente piense en términos de *consecuencias imaginadas*. Pero entonces parece que se aparta cada vez más de las ventajas iniciales del utilitarismo. Pues, primero, el supuesto cálculo de la utilidad de que la gente piense en términos de consecuencias imaginadas —más bien que examinando los casos particulares, o que examinando las consecuencias de reglas más específicas, o que considerando la moralidad local como un asunto de su incumbencia, o que muchas otras posibilidades—, ese cálculo empieza a tener cada vez más aspecto de ser engañoso en sus pretensiones. ¿Cómo podría saber el utilitarista cuáles podrían ser las consecuencias de esas diversas prácticas? En segundo lugar, si que sabe una cosa por lo menos: que cuanto más general sea la práctica a la que se aplica el cálculo utilitarista, tantos más serán los casos en los que el cálculo particular para ese caso habría conducido a un resultado distinto, y con ello, tanto mayor también la inutilidad táctica que está permitiendo en su persecución de la uti-



lidad estratégica. Y en vistas de lo problemática que resulta esa utilidad estratégica, no tiene más remedio que preocuparle lo siguiente: una de las motivaciones del utilitarismo era, después de todo, un obstinado imperativo de pensar en términos de consecuencias calculables y no atenerse a la tradición, a las prácticas recibidas, etc.

Cuanto más considera uno el utilitarismo de reglas, tanto más urgentes resultan las cuestiones de esta clase. Volviendo una vez más al modelo del recibo del gas, recordemos que lo que principalmente hacía sensata esa práctica uniforme eran los costos que suponía el interferirla. Lo análogo de esto en la deliberación moral ordinaria es la inutilidad de calcular las consecuencias particulares. Pero el efecto del argumento queda cancelado si consideramos un caso en el que el cálculo particular *ha sido hecho ya*; y esto es así en los casos moralmente problemáticos que representaban la segunda clase de dificultades que se suponían quedaban solventadas por el utilitarismo de reglas. Si el cálculo se ha hecho ya, y ha resultado que las consecuencias de infringir la regla son mejores que las de respetarla, entonces las consideraciones sobre la inutilidad del cálculo no pueden cambiar este resultado. Y, desde luego, es difícil ver que, para un utilitarista coherente, pudiera haber algo, sea lo que ello fuere, capaz de cambiar este resultado. Cualquiera que sea la utilidad general de contar con una cierta regla, si uno ha alcanzado de hecho un punto desde el que ve que la utilidad de infringir la regla en una determinada ocasión es mayor que la de respetarla, entonces no cabe duda de que sería totalmente irracional no infringirla.

Ha habido, desde luego, utilitaristas duros que han sacado esa conclusión, como es el caso de J. J. C. Smart. Si al utilitarismo se le puede hacer funcionar en absoluto, entonces para mí está fuera de duda de que la actitud correcta ha de ser la de estos autores: el utilitarismo es una doctrina especial, que no tiene necesariamente que coincidir con las ideas morales contemporáneas de Occidente en todos los aspectos, y, siendo así, uno ha de esperar que tenga lo que muy bien podrían parecer con-

clusiones poco gratas. Contrariamente a esto, un rasgo de la moderna teoría utilitarista es que es una teoría sorprendentemente conformista. Bentham y Mill consideraron el principio de la mayor felicidad como un instrumento de crítica, y pensaban que, valiéndose de él, podrían mostrar que buena parte de las creencias morales de la época victoriana eran erróneas e irracionales, como por supuesto lo eran. Pero, con excepción de las bien delimitadas áreas de la reforma sexual y de la reforma penal, cuyo tratamiento es el heredado de Bentham y Mill, la teoría utilitarista moderna tiende a gastar más esfuerzo en reconciliar el utilitarismo con las creencias morales existentes que en rechazar esas creencias apoyándose en el utilitarismo. Un autor reciente, por ejemplo, se ha tomado un honesto y gran esfuerzo en mostrar que las ejecuciones públicas no podrían justificarse, como podría parecer, con argumentos utilitaristas. Nuestro autor se queda con algunas francas dudas; pero éstas son dudas sobre la aplicación y la formulación del utilitarismo, y no, como deberían ser, dudas sobre si no podrían ser reintroducidas las ejecuciones públicas. Este es un caso absurdo. Pero en términos más generales, todas las cualidades humanas que son consideradas valiosas y que, sin embargo, se resisten a un tratamiento utilitarista, como son la intransigente pasión por la justicia, ciertas clases de coraje, la espontaneidad, la disposición a oponerse a cosas tales como los experimentos útiles sobre pacientes ancianos o las bombas de napalm sobre ciertas gentes para asegurar (se dice) la felicidad de más gente, generan con frecuencia en los teóricos del utilitarismo el deseo de acomodar el utilitarismo a esos valores, y no la condena de tales valores como un legado irracional de la era preutilitarista. Esto es sin duda un tributo a la honestidad e imaginación de esos utilitaristas, pero no a su consistencia o a su utilitarismo.

El utilitarismo de reglas, como la empresa que es de tratar de atenerse a algo distintivamente utilitarista, procurando al mismo tiempo quitarle sus puntas más hirientes, me parece un fracaso. Ese terreno intermedio no es habitable lógicamente. Frente a él, uno podría,

por un lado, adoptar la línea a la que ya me he referido, la de Smart y otros, y seguir con el utilitarismo de actos modificado solamente dentro de los límites reconocibles del modelo del recibo del gas. Esto es, por lo menos, coherente con los equivocados fines del utilitarismo, y el hecho de que arroje algunos resultados morales particulares distintivos y (para muchos) indigeribles no debería ser motivo de sorpresa. Si, por otro lado, se abandona este territorio, y se empieza a aplicar el principio utilitarista a prácticas y a hábitos de pensamiento cada vez más generales, es muy improbable que lo que que quede tenga algún contenido que pueda considerarse distintivamente utilitarista. Esta capacidad del utilitarismo para aniquilarse a sí mismo, una vez abandonado el nivel básico, puede quedar ilustrada con un breve argumento, con el que voy a terminar. Sus premisas empíricas quizás no estén fuera de toda duda. Pero sí que son por lo menos tan plausibles como las que los utilitaristas suelen utilizar en estas materias.

Uno de los efectos perturbadores cuando la gente se convierte en utilitaristas activos y conscientes es que tienden a degradar el patrón de medida moral: comienza a operar una ley de Gresham por la que las malas acciones de los hombres malos obliga a hombres mejores a hacer cosas que, en mejores circunstancias, serían también malas. La razón de esto es simple: un utilitarista está siempre justificado a hacer la cosa menos mala que sea necesaria para evitar la cosa peor, cosa peor que, si no lo impidiera la menos mala, ocurriría en las circunstancias dadas (incluyendo, por supuesto, la cosa peor que cualquier otro podría hacer) y lo que de esta forma está justificado a hacer puede ser algo que, considerado en sí mismo, es bastante sucio. El acto de evitación forma parte de las concepciones utilitaristas, y ciertas nociones de responsabilidad negativa (el que uno sea responsable de lo que deja de evitar tanto como de lo que hace) son, por lo mismo, también características del utilitarismo. Si esto es así, es empíricamente probable que pueda esperarse una escalada de la actividad de evitación; y las consecuencias totales de esto, *medidas en los propios*

*estándares utilitaristas*, serán peores que si nunca se hubiera iniciado esa escalada.

Sin embargo, el utilitarista que está inmerso en el sistema no puede hacer nada para evitar esto; tiene que pensar en términos de las consecuencias fácticas en curso, y ningún amago de principio podrá hacer ahora nada (o por lo menos, nada útil) en el terreno de esas consecuencias fácticas en curso —desde donde está no hay forma de dirigir una incursión a un terreno moralmente más alto. Volviendo, sin embargo, sobre sí mismo, puede considerar cómo los objetivos utilitaristas podrían haber sido realizados mejor de lo que lo han sido en un mundo de utilitaristas mezclados con villanos. Y sin duda que hubieran sido mejor realizados si no hubiera habido villanos: pero esto es ciertamente utópico. Lo más prometedor parece que es un estado de cosas en que un número suficiente de personas se resista a proseguir esta carrera de envilecimiento: que se resistan, por ejemplo, fijando una serie de cosas que no puede entrar en consideración hacer, o que no están dispuestos ellos a hacer o que no pueden soportar que se hagan, cualesquiera sean las cosas que los otros hacen o pueden hacer. Existe un límite a sus actividades de evitación. Y para ello parece que suficiente gente, durante un tiempo suficiente, tiene que estar dispuesta a obstinarse en hacer varias cosas, cualesquiera que sean las consecuencias. Esto significa que suficiente gente, durante un tiempo suficiente, no han de pensar como utilitaristas; sin duda alguna, han de pensar como no utilitaristas. Y no podrán preservar en la trastienda de sus mentes la actitud utilitarista en coexistencia con la requerida obstinación moral. Pues tienen que ser capaces de resistir la tentación utilitaria en las circunstancias más difíciles, es decir, cuando es obvio que de resistirla se van a seguir muchos males, y para esto su no-utilitarismo tiene que estar profundamente enraizado.

Algunos utilitaristas han llegado, aunque no exactamente por las mismas razones, a algo parecido a esta conclusión y han pensado que lo que con ello se demostraba es que de la verdad del utilitarismo podría estar

al tanto una élite responsable, pero que no debería estar excesivamente divulgada entre las masas. Esta propuesta, tanto desde un punto de vista personal como desde un punto de vista social, es una propuesta sin perspectivas. Desde un punto de vista personal, porque el estado mental que se atribuye al utilitarista reflexivo, y la actitud hacia otro que ese estado implica, sólo podrían ser honestamente mantenidas, si es que pueden serlo, por un hombre muy inocente (tal vez Sidgwick lo era), y ningún hombre reflexivo de nuestra época puede ser ese inocente. Desde un punto de vista social, porque las instituciones educativas y otras instituciones que se requerirían para encarnar esa idea tendrían que ser muy diferentes de todo lo que ahora podemos esperar o tolerar, o de todo lo que el propio utilitarismo podría desear.

Si todo esto es verdad, entonces el mundo que el utilitarista reflexivo tiene finalmente que escoger, como el mundo que más probablemente puede arrojar los resultados que el utilitarista desea, es un mundo en el que la ley de Gresham no rige porque un número suficiente de personas durante un tiempo suficiente se muestran profundamente recalcitrantes a pensar en términos utilitaristas. No es posible que esta disposición pudiera coexistir con la creencia en el utilitarismo; ni tampoco parece aceptable o socialmente posible que la mayoría tuviera que tener esta disposición, mientras que otros, la élite utilitarista, creyeran en el utilitarismo. Todo lo que queda es que el mundo que satisfaría las aspiraciones del utilitarista sería un mundo del que estuviera totalmente ausente la creencia en el utilitarismo como una doctrina moral omnicomprendensiva, excepto tal vez como una excentricidad menor y sin trascendencia.

De este modo, si el utilitarismo es verdadero, y si también son verdaderas algunas premisas empíricas bastante plausibles, entonces es mejor que la gente no crea en el utilitarismo. Si por otro lado, el utilitarismo es falso, entonces por supuesto que es mejor que la gente no crea en él. Así pues, de cualquier forma es mejor que la gente no crea en él.